August 10, 2009

WHY THE THEORY R IS SPECIAL

ALBERT VISSER

ABSTRACT. Is it possible to give coordinate-free characterizations of salient theories? Such characterizations would always involve some notion of sameness of theories: we want to describe a theory modulo a notion of sameness, without having to give an axiomatization in a specific language. Such a characterization could, e.g., be a first order formula in the language of partial preorderings that describes uniquely a degree in a particular structure of degrees of interpretability. Our theory would be contained in this degree. There are very few examples currently known along these lines, except some rather trivial ones.

In this paper we provide a non-trivial characterization of Tarski-Mostowski-Robinson's theory R. The characterization is in terms of the double degree structure of RE degrees of local and global interpretability. Consider the RE degrees of global interpretability that are in the minimal RE degree of local interpretability. These are the global degrees of the RE locally finitely satisfiable theories. We show that these degrees have a maximum and that R is in that maximum. In more mundane terms: an RE theory is locally finite iff it is globally interpretable in R.

Dedicated to Harvey Friedman on the occasion of his 60th birthday.

1. Introduction

Wouldn't it be nice if we could characterize salient theories like Robinson's Arithmetic and Peano Arithmetic in a coordinate-free way, independent of particular choices of signature and axiomatization? A moment's reflection shows that such a characterization would only be possible modulo some notion of sameness of theories. For example, one could imagine that a certain RE degree of interpretability was characterized by a first-order formula A over the partial preorder of degrees of interpretability of RE theories. If our salient theory were in that degree, our formula A would be the coordinate-free characterization we are looking for. It would be characterized modulo mutual interpretability.

Regrettably, we have very few examples of such characterizations and the ones we have are rather trivial. Here are three examples.

- The theory EQ of pure equality is the initial element of the category of direct interpretations (without parameters). (An interpretation is *direct* if it is identity preserving and unrelativized.) The theory EQ is thus characterized modulo synonymy (aka. definitional equivalence).
- Consider the lattice of degrees of local interpretability of theories, where we impose no restriction on the signature nor on the complexity of the theory. The intended notion of interpretation is more-dimensional interpretation with parameters. This structure is studied by Mycielski, Pudlák and Stern in [MPS90]. They call these degrees: *chapters*. The maximum of the

2000 *Mathematics Subject Classification.* 03A05, 03B30, 03F25,
Key words and phrases. interpretability, local interpretability, finite satisfiability.

structure is the degree of inconsistent theories. There is also an element directly below this maximum: the maximal degree of consistent theories. The following two salient theories are in the degree: $\mathsf{Th}(\mathbb{N})$ and $\mathsf{Th}_{\Pi^0_1}(\mathbb{N})$ (the theory axiomatized by the true Π^0_1-sentences).[1]

- Consider the degree structure of degrees of one-dimensional interpretability with parameters. Let $\uparrow\mathbf{n}$, for $n = 1, 2, \ldots$, be EQ, the theory of pure equality, plus the sentence $\exists x_0, \ldots, x_{n-1} \bigwedge_{i<j<n} x_i \neq x_j$. Let $\uparrow\infty$ be the union of the $\uparrow\mathbf{n}$. Our degree structure yields a partial ordering of the following form: we first have $\omega + 1$ and above that something else. The theory $\uparrow\mathbf{n}$ is in the $(n-1)$th degree from below and $\uparrow\infty$ is in the ωth degree. So each of these theories is characterized modulo mutual one-dimensional interpretability with parameters. (See Theorem 3.3 of this paper.)

In the present paper, we produce a less trivial example of a characterization. The structure in which the characterization is given is the double degree structure of local and global interpretability for RE theories. The notion of interpretation involved is: more-dimensional, piecewise interpretation with parameters.[2] We will show that the theory R is the maximum of the global degrees that are in the minimal local degree.[3] As we will see this means that an RE theory is locally finitely satisfiable iff it is globally interpretable in R.

The theory R was introduced by Tarski, Mostowski and Robinson in their book [TMR53]. It is a very weak theory that is *essentially undecidable.* This means that every consistent RE extension of the theory is undecidable. It was observed by Cobham that one still has an essentially undecidable theory if one drops the axiom $\mathsf{R6}$ (given below), obtaining the theory R_0. See [Vau62] and [JS83]. Cobham has shown that R has a stronger property than essential undecidability. Consider any RE theory T. Suppose we have translation α of the arithmetical language into the language of T. Suppose T is consistent with R_0^α. Then, T is undecidable.[4] For the proof of a closely related result, see Vaught's paper [Vau62]. In fact one can show that, if T is consistent with R_0^α, then there is a finitely axiomatized extension A of R_0 and a translation β, such that T is consistent with A^β.

2. Theories and Interpretations

In this section, we fix some basic concepts and notations. The reader is advised to go over it lightly, returning just when a notation or notion is not clear.

We work with RE theories in one-sorted first order predicate logic of finite signature. These theories have *officially* a relational signature. Unofficially, we use function symbols, but these can be eliminated using a well-known unwinding procedure. Every sort has identity.

Our most general notion of interpretation is *piecewise, more-dimensional, relative interpretation with parameters, where identity is not necessarily translated as*

---

[1]The result also holds if we restrict ourselves to one-dimensional interpretations and/or parameter-free interpretations.

[2]We will explain *piecewise* in the paper.

[3]We will prove a number of related results where we vary the notion of interpretation used in defining the degree structure.

[4]Cobham's proof remains unpublished, but, using the clues provided in [Vau62], it is not hard to find a proof.

identity. We will first set up the machinery for the case of *more-dimensional, relative interpretation with parameters, where identity is not necessarily translated as identity*. Then, we will extend the framework to piecewiseness. Even if the basic idea behind the various kinds of interpretation is rather obvious, some care is needed to get the definitions right, mainly because some careful management of the use of variables is necessary.

2.1. **Translations.** To define an interpretation, we first need the notion of *translation*. We first define the notion of *many-dimensional translation with parameters*.

Let Σ and Ξ be finite signatures for first-order predicate logic. We fix a sequence containing all variables $u_0, u_1, \ldots$ for the signature Σ, and we fix three disjoint sequences of variables $v_0, v_1, \ldots$, and $w_0, w_1, \ldots$, and $z_0, z_1, \ldots$ for the signature Ξ.

A *relative translation* $\tau : \Sigma \to \Xi$ is given by a quintuple $\langle n, m, \pi, \delta, F\rangle$. Here n and m are natural numbers. The number n is the *dimension* of the interpretations and the number m gives us the number of parameters. The formula π is Ξ-formula with free variables among $w_0, \ldots, w_{m-1}$. This formula gives a constraint on the possible parameters. The formula δ is a Ξ-formula with bound variables among the $z_0, z_1, \ldots$ and free variables among $v_0, \ldots, v_{n-1}, w_0, \ldots, w_{m-1}$. It defines the domain of the interpretation. The mapping F associates to each relation-symbol R of Σ a Ξ-formula $F(R)$. Let the arity of R be k. We demand that the bound variables of $F(R)$ are among the z_i, and that $F(R)$ has at most the variables $v_0, \ldots, v_{kn-1}, w_0, \ldots, w_{m-1}$ free. We will write $\vec{v}_i$ for the block of variables $v_{ni}, \ldots, v_{n(i+1)-1}$ and $\vec{w}$ for $w_0, \ldots, w_{m-1}$. So, $F(R)$ will have at most $\vec{v}_0, \ldots, \vec{v}_{k-1}, \vec{w}$ free. We translate Σ-formulas to Ξ-formulas as follows:

- $(R(u_{j_0}, \cdots, u_{j_{k-1}}))^\tau := F(R)(\vec{v}_{j_0}, \cdots, \vec{v}_{j_{k-1}}, \vec{w})$;
  the formula $F(R)(\vec{v}_{j_0}, \ldots)$ is the result of simultaneous substitution of v_{nj_i+s} for v_{ni+s}, where $0 \leq i < k$ and $0 \leq s < n$;
- $(\cdot)^\tau$ commutes with the propositional connectives;
- $(\forall u_k\, A)^\tau := \forall \vec{v}_k (\delta(\vec{v}_k, \vec{w}) \to A^\tau)$;
- $(\exists u_k\, A)^\tau := \exists \vec{v}_k\, (\delta(\vec{v}_k, \vec{w}) \wedge A^\tau)$.

We have introduced translations with careful variable management, showing how this can be done. In practice we want to be sloppy. We use e.g. x and x_i as metavariables ranging over the u_j (and we will overload the use of e.g. z in the same way). If x stands for u_i, we write $\vec{x}$ for $\vec{v}_i$. If $\vec{x}$ stands for $x_0, \ldots, , x_{k-1}$, we will use $\vec{\vec{x}}$ for $\vec{x}_0, \ldots, \vec{x}_{k-1}$. Etc.

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

- We write δ_τ for δ, etc.
- We write R_τ for $F_\tau(R)$.
- We write $\vec{\vec{x}} : \delta_{\vec{w}}$ for: $\delta(\vec{x}_0, \vec{w}) \wedge \ldots \wedge \delta(\vec{x}_{k-1}, \vec{w})$.
- We write $\forall \vec{\vec{x}}{:}\delta_{\vec{w}}\, A$ for: $\forall \vec{x}_0 \ldots \forall \vec{x}_{k-1}\, (\vec{\vec{x}}{:}\delta_{\vec{w}} \to A)$. Similarly for the existential case.

We can define the identity translation and composition of translations in the obvious way.

Consider $\tau = \langle n, m, \pi, \delta, F\rangle$. The translation τ is one-dimensional if $n = 1$. In this case we will write: $\tau = \langle m, \pi, \delta, F\rangle$. The translation τ is parameter-free if $m = 0$ and $\pi = \top$. In this case, we will write: $\tau = \langle n, \delta, F\rangle$. If τ is one-dimensional and parameter-free, we write: $\tau = \langle \delta, F\rangle$.

2.2. **Interpretations and Interpretability.** A translation τ supports a *relative interpretation* of a theory U in a theory V, if, $V \vdash \exists \vec{w}\ \pi_\tau \vec{w}$ and, for all sentences A of the language of U, we have $U \vdash A \Rightarrow V \vdash \forall \vec{w}\ (\pi_\tau \rightarrow A^\tau)$. Note that this automatically takes care of the theory of identity. Moreover, it follows that $V \vdash \forall \vec{w}\ (\pi_\tau \vec{w} \rightarrow \exists \vec{v}\ \delta_\tau \vec{v}\vec{w})$. Thus, an interpretation has the form: $K = \langle U, \tau, V\rangle$. We can define the identity interpretation and composition of interpretations in the obvious way.

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

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

We say that *a theory* V *locally interprets a theory* U if, for any finite subtheory U_0 of U, we have $V \rhd U_0$. We write $V \rhd_{\mathsf{loc}} U$ for: V locally interprets U. It is easily see that both $\lhd$ and $\lhd_{\mathsf{loc}}$ are preorderings.

2.3. **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.

A translation τ now has the form $\langle \ell, \nu, m, \pi, \delta, F\rangle$, where:

- ℓ is a natural number that stands for the set of pieces $0, \ldots, \ell - 1$;
- ν is a function that assigns to each piece j a dimension ν^j;
- m is again the number of parameters, and π a constraint on the parameters;
- δ is a function from pieces j to domains δ^j of dimension ν^j;
- F is now a function that sends a pair P, f to an appropriate formula. Here f assigns to each argument place of P a piece,

To make it all work smoothly we again enumerate the variables of Σ by $u_0, u_1, \ldots$ We fix $\ell + 2$ disjoint sequences of variables for the signature Ξ, to wit $v^0_0, v^0_1 \ldots$, and $\ldots$ and $v^{\ell-1}_0, v^{\ell-1}_1, \ldots$, and $w_0, w_1, \ldots$ and $z_0, z_1, \ldots$ We write $\vec{v}^{\,j}_i$ for $v^j_{\nu^j i}, \ldots, v^j_{\nu^j (i+1)-1}$ and $\vec{w}$ for $w_0, \ldots, w_{m-1}$. We demand:

- δ^j has bound variables among the $z_0, z_1, \ldots$ and free variables among $\vec{v}^{\,j}_0, \vec{w}$.
- Suppose P has arity k. $F(P, f)$ has bound variables among $z_0, z_1, \ldots$ It has free variables among $\vec{v}^{\,f0}_0, \vec{v}^{\,f1}_1, \ldots, \vec{v}^{\,f(k-1)}_{k-1}, \vec{w}$.

We translate Σ-formulas to Ξ-formulas as follows. The basic form of translation is $A^{\tau,g}$, where g is a function from the indices of the free variables of A to pieces.

- $(R(u_{j_0}, \cdots, u_{j_{k-1}}))^{\tau,g} := F(R, f)(\vec{v}^{\,gj_0}_{j_0}, \cdots, \vec{v}^{\,gj_{k-1}}_{j_{k-1}}, \vec{w})$, where $fs := g(j_s)$; the formula $F(R, f)(\vec{v}^{\,gj_0}_{j_0}, \ldots)$ is the result of simultaneous substitution of $v^{fi}_{\nu^{fi} j_i + s}$ for $v^{fi}_{\nu^{fi} i + s}$, where $0 \leq i < k$ and $0 \leq s \leq \nu^{fi} - 1$.
- $(A \wedge B)^{\tau,g} := A^{\tau,g\restriction \mathsf{fv}(A)} \wedge B^{\tau,g\restriction \mathsf{fv}(B)}$;
  similarly for the other propositional connectives;
- $(\forall u_k\, A)^{\tau,g} := \bigwedge_{j<\ell} \forall \vec{v}^{\,j}_k\ (\delta^j(\vec{v}^{\,j}_k, \vec{w}) \rightarrow A^{\tau,g[k:j]})$, where $g[k:j]$ is the result of setting g at k to j;

- $(\exists u_k\, A)^{\tau,g} := \bigvee_{j<\ell} \exists \vec{v}^{\,j}_k\ (\delta^j(\vec{v}^{\,j}_k, \vec{w}) \wedge A^{\tau,g[k:j]})$.

A translation τ is called *parameter-free* if $m_\tau = 0$ and $\pi = \top$, *one-dimensional* if $\nu_\tau j = 1$, for all $j < \ell_\tau$, *one-piece* iff $\ell_\tau = 1$. Similarly, for interpretations.

The rest of the development is the same as before.

Example 2.1. We show that $\uparrow\mathbf{0} := \mathsf{EQ}$, the theory of pure identity interprets $\uparrow\mathbf{n} := \mathsf{EQ} + \exists u_0, \ldots, u_{\ell-1}\ \bigwedge_{i<j<\ell} u_i \neq u_j$. We take τ with:

- $\ell_\tau := \ell$,
- $\nu_{\tau,j} = 1$,
- $m = 0$,
- $\pi_\tau := \top$,
- $\delta^{\tau,j} := (v^j_0 = v^j_0)$,
- $F_\tau(=, ij) := (v^i_0 = v^i_1)$, if $i = j$, and $F_\tau(=, ij) := \bot$, otherwise.

We can easily see that this construction does indeed yield the desired interpretation. It follows that $\uparrow\mathbf{0} \rhd_{\mathsf{loc}} \uparrow\infty$. ❑

If the target theory V proves that we have at least two elements, we can always replace a piecewise interpretation by a more-dimensional one. Here is how this works. Suppose there are ℓ pieces and that k is the maximum of the dimensions of the pieces. Our new one-piece interpretation will be $k+\ell$-dimensional. We represent an object $v^j_0, \ldots v^j_{\nu^j-1}$ of δ^j as $v_0, \ldots v_{k+\ell-1}$. Here we demand that $\delta^j(v_0, \ldots, v_{\nu^j-1})$. The variables $v_{\nu^j}, \ldots, v_{m-1}$ are unconstrained: they serve as padding to get the right length. The variables $v_m, \ldots v_{\ell-1}$ satisfy $v_m = \ldots = v_{m+\nu_j-1} \neq v_{m+\nu_j} = \ldots = v_{k+\ell-1}$. They serve to keep the sequences from different pieces disjoint. From each sequence we can read off from which piece it comes. To find out whether $v_0, \ldots, v_{k+\ell-1}$ is equal to $v_{k+\ell}, \ldots, v_{2(k+\ell)-1}$, we first extract j_0 and j_1 from the end-strings and then check whether $F(=, j_0 j_1)(v_0, \ldots, v_{\nu^{j_0}-1}, v_{k+\ell}, \ldots, v_{k+\ell+\nu^{j_1}-1})$.

The treatment of the other predicates is similar.

3. Double Degree Structures

The double degree structure $\mathcal{D}_{ijk}$ consists of the RE theories plus two partial preorders $\lhd$ and $\lhd_{\mathsf{loc}}$. Here the interpretations considered are (possibly) more-dimensional if $i = 1$ and 1-dimensional if $i = 0$, (possibly) piecewise if $j = 1$ and one-piece if $j = 0$, (possibly) with parameters if $k = 1$ and parameter-free if $k = 0$. Note that $\lhd$ is a subordering of $\lhd_{\mathsf{loc}}$.

The basic degree structures we work with are given as pairs of partial preorders. However, when convenient, we will flexibly switch to talk about degrees where the induced equivalence relations are divided out. Note that if we think of the degree structures in this last way, there is a projection functor π mapping the degrees of global interpretability to the degrees of local interpretablity that is part of the double degree structure. We will use $[T]$ for the global degree of T and $[T]_{\mathsf{loc}}$ for the local degree of T.

Let the lowest degree of global interpretability be $\perp\!\!\!\perp$ and let the lowest degree of local interpretability be $\perp\!\!\!\perp_{\mathsf{loc}}$. So, $\perp\!\!\!\perp = [\mathsf{EQ}]$ and $\perp\!\!\!\perp = [\mathsf{EQ}]_{\mathsf{loc}}$.

We provide a convenient characterization of $\perp\!\!\!\perp$ and $\perp\!\!\!\perp_{\mathsf{loc}}$ in $\mathcal{D}_{ijk}$ with $j = 1$. So only the presence of piecewise interpretations is essential for the result. A theory T is *finitely satisfiable* if T has a finite model. A theory T is *locally finitely satisfiable* iff every finitely axiomatized subtheory of T has a finite model.

Theorem 3.1. *We work in $\mathcal{D}_{ijk}$ with $j = 1$. (i) The RE theory T is in $\perp\!\!\!\perp$ iff T is finitely satisfiable. (ii) The RE theory T is in $\perp\!\!\!\perp_{\mathsf{loc}}$ iff T is locally finitely satisfiable.*

Proof. Suppose T is in $\perp\!\!\!\perp$. So, for some K, we have: $K : \mathsf{EQ} \rhd T_0$. Let $\mathcal{M}$ be any finite model of EQ. Clearly, the interpretation K gives a finite 'internal' model $K(\mathcal{M})$ of $\mathcal{M}$.

Conversely, suppose T is finitely satisfiable. Let $\mathcal{M}$ be a finite model of T. We may assume that the elements of $\mathcal{M}$ are $0, \ldots, m-1$. We now construct an interpretation that 'describes' $\mathcal{M}$ as follows. We take τ with:

- $\ell := m$,
- $\nu_j = 1$,
- $m = 0$,
- $\pi := \top$,
- $\delta^j := (v_0^j = v_0^j)$,
- $F(P, j_0 \ldots j_{k-1}) := \top$, if $P_{\mathcal{M}}(j_0, \ldots, j_{k-1})$, and $F(P, j_0 \ldots j_{k-1}) := \perp$, otherwise.

Part (ii) is immediate form (i). ❑

What happens when we do not have piecewise interpretations. Let us first consider the cases of $\mathcal{D}_{100}$ and $\mathcal{D}_{101}$. We have the following theorem.

Theorem 3.2. *Consider $\mathcal{D}_{100}$ or $\mathcal{D}_{101}$.*

i. Both $\perp\!\!\!\perp$ and $\perp\!\!\!\perp_{\mathsf{loc}}$ consist of precisely the theories that have a one-element model.
ii. The degree $\perp\!\!\!\perp$ has a unique immediate successor $\perp\!\!\!\perp^+$, that consists of the finitely satisfiable theories for which every model has at least two elements.
iii. The degree $\perp\!\!\!\perp_{\mathsf{loc}}$ has an a unique immediate successor $\perp\!\!\!\perp^+_{\mathsf{loc}}$, that consists of the locally finitely satisfiable theories, for which every model has at least two elements.

Note that it follows that the (locally) finitely satisfiable theories can be characterized in terms of the degree structure.

Proof. Let $\uparrow\mathbf{1} := \mathsf{EQ}$ and $\uparrow\mathbf{2} := (\mathsf{EQ} + \exists x, y\; x \neq y)$.

Ad (i): It is easy to see that the elements of $\perp\!\!\!\perp$ are precisely those with a one-element model. Consider any T such that $\mathsf{EQ} \rhd_{\mathsf{loc}} T$. It follows that every finite subtheory of T has a one-element model. Since the signature is finite, there are only finitely many such one-element models. So, there must be one model that satisfies arbitrarily large finite subtheories of T and, hence, T itself.

Ad (ii): We take $\perp\!\!\!\perp^+ := [\uparrow\mathbf{2}]$. We clearly have $\uparrow\mathbf{2} \underset{\neq}{\unrhd} \uparrow\mathbf{1}$.

Suppose T is finitely satisfiable and every model of T has at least two elements. It is immediate that $T \vdash \uparrow\mathbf{2}$, and hence $T \rhd \uparrow\mathbf{2}$. By the considerations at the end of Subsection 2.3, we can simulate piecewise interpretations as soon as we have two elements available in the interpreting theory. Thus, $\uparrow\mathbf{2} \rhd T$. So if T is finitely satisfiable, then $T \equiv \uparrow\mathbf{2}$.

For the converse, suppose $T \equiv \uparrow\mathbf{2}$. It is easy to see that T must be finitely satisfiable. Moreover $T \rhd \uparrow\mathbf{2}$, implies $T \vdash \uparrow\mathbf{2}$, hence every model of T has at least two elements.

Finally consider any RE theory W. If W has a one-element model then $W \equiv \uparrow\mathbf{1}$, if not then $W \rhd \uparrow\mathbf{2}$.

Ad (iii): This is immediate from (ii). ❑

Finally, we consider $\mathcal{D}_{001}$. Thus we will leave the case of $\mathcal{D}_{000}$ open. We remind the reader that, for $n > 0$, $\uparrow\mathbf{n} := (\mathsf{EQ} + \exists x_0, \ldots, x_{n-1} \bigwedge_{i<j<n} x_i \neq x_j)$, and $\uparrow\infty$ is the union of the $\uparrow\mathbf{n}$.

Theorem 3.3. *In $\mathcal{D}_{001}$ the situation is as follows.*

i. *A theory T is finitely satisfiable iff $T \equiv_{(\mathsf{loc})} \uparrow\mathbf{n}$, for some n. Here n is the minimal size of a model satisfying T.*
ii. $\uparrow\mathbf{1} \trianglelefteq\!\!\!\!\!/_{(\mathsf{loc})} \uparrow\mathbf{2} \ldots \trianglelefteq\!\!\!\!\!/_{(\mathsf{loc})} \uparrow\infty$.
iii. *For any T, $T \equiv_{(\mathsf{loc})} \uparrow\mathbf{n}$ for some n, or $T \rhd_{(\mathsf{loc})} \uparrow\infty$.*
iv. *A theory T is locally finitely satisfiable iff $T \lhd_{\mathsf{loc}} \uparrow\infty$.*
v. *$T \equiv_{\mathsf{loc}} \uparrow\infty$ iff T is locally finitely satisfiable and has only infinite models, i.o.w. iff T is locally finitely satisfiable but not finitely satisfiable.*

Note that the degrees $[\uparrow\mathbf{n}]$, $[\uparrow\infty]$, $[\uparrow\mathbf{n}]_{\mathsf{loc}}$ and $[\uparrow\infty]_{\mathsf{loc}}$ are all determined in terms of the degree structure.

Proof. We skip the easy proof. Note that we need the parameters to 'describe' a model of n or less than n elements in $\uparrow\mathbf{n}$. ❑

The theory R, due to Tarski, Mostowski and Robinson ([TMR53]) is locally finitely satisfiable and not finitely satisfiable. The main result of our paper is that, for any locally finitely satisfiable theory T is interpretable in R via a one-dimensional, one-piece, parameter-free, identity preserving interpretation. As a direct consequence of the above results, we find:

- In $\mathcal{D}_{ijk}$ with $j = 1$, R is in the $\lhd$-maximum of $\perp\!\!\!\perp_{\mathsf{loc}}$,
- In $\mathcal{D}_{100}$ and $\mathcal{D}_{101}$, R is in the $\lhd$-maximum of $\perp\!\!\!\perp^{+}_{\mathsf{loc}}$.
- In $\mathcal{D}_{001}$, R is in the $\lhd$-maximum of $[\uparrow\infty]_{\mathsf{loc}}$.

In all cases the degree [R] is first order definable in the double degree structure.

4. The Theories R, Q$^-$ and Q

In this section, we briefly introduce three theories of number theory R, Q$^-$ and Q. The theories Q$^-$ and Q will each play a role in the proof of our main theorem. We consider the signature with constant and function symbols 0, S, + and $\cdot$. We define $\underline{0} := 0$, $\underline{n+1} := \mathsf{S}\underline{n}$, and $x \leq y :\leftrightarrow \exists z\, z + x = y$. We consider the following axioms.

R1. $\vdash \mathsf{S}\underline{n} = \underline{n+1}$
R2. $\vdash \underline{m} + \underline{n} = \underline{m+n}$
R3. $\vdash \underline{m} \cdot \underline{n} = \underline{m \cdot n}$
R4. $\vdash \underline{m} \neq \underline{n}$, for $m \neq n$
R5. $\vdash x \leq \underline{n} \to \bigvee_{i \leq n} x = \underline{i}$
R6. $\vdash\vdash x \leq \underline{n} \vee \underline{n} \leq x$
R7. $\vdash\ \leq$ is a linear ordering
R8. $\vdash x < y \to \mathsf{S}x \leq y$
R9. $\vdash x \leq y \to \exists z{\leq}y\; z + x = y$

The theory R$_0$ is axiomatized by R1,2,3,4,5. The theory R is axiomatized by R1,2,3,4,5,6.[5] The theory R* is axiomatized by R1,2,3,4,5,7,8,9.

[5]The original version of R does not have S, but a constant 1. However it is definitionally equivalent with our version: The original version can be recovered from ours by translating 1 to S0. Our version can be recovered from the original one by translating Sx to $x + 1$.

It was observed by Cobham that one still has an essentially undecidable theory if one drops the axiom R6 (given below), obtaining the theory R_0. See [Vau62] and [JS83]. In fact R is interpretable in R_0. In [Vis09a], it is shown that $\mathsf{R}^\star$ is interpretable in R_0, and hence in R.

Robinson's Arithmetic Q was introduced in [TMR53]. Using Solovay's method of shortening cuts (see [Solle]), one can show that Q interprets seemingly much stronger theories like $I\Delta_0 + \Omega_1$. See [Nel86] and [HP91]. Here are the axioms of Q.

Q1. $\vdash \mathsf{S}x = \mathsf{S}y \to x = y$
Q2. $\vdash 0 \neq \mathsf{S}x$
Q3. $\vdash x = 0 \vee \exists y\, x = \mathsf{S}y$
Q4. $\vdash x + 0 = x$
Q5. $\vdash x + \mathsf{S}y = \mathsf{S}(x + y)$
Q6. $\vdash x \times 0 = 0$
Q7. $\vdash x \times \mathsf{S}y = x \times y + x$

The theory Q^- is due to Andrzej Grzegorczyk. It is a prima facie weakening of Q in which addition and multiplication are partial. Thus, in stead of $+$ and $\times$ we have ternary relation symbols A and M. We define $\leq$ by $x \leq y :\leftrightarrow \exists z\, \mathsf{A}zxy$. The theory is axiomatized as follows.

G1. $\vdash \mathsf{S}x = \mathsf{S}y \to x = y$,
G2. $\vdash 0 \neq \mathsf{S}x$,
G3. $\vdash x = 0 \vee \exists y\, x = \mathsf{S}y$,
G4. $\vdash (\mathsf{A}xyz_0 \wedge \mathsf{A}xyz_1) \to z_0 = z_1$
G5. $\vdash \mathsf{A}x0x$
G6. $\vdash \mathsf{A}xyz \to \mathsf{A}x(\mathsf{S}y)(\mathsf{S}z)$
G7. $\vdash (\mathsf{M}xyz_0 \wedge \mathsf{M}xyz_1) \to z_0 = z_1$
G8. $\vdash \mathsf{M}x00$
G9. $\vdash (\mathsf{M}xyz \wedge \mathsf{A}zxw) \to \mathsf{M}x(\mathsf{S}y)w$

The theory Q is interpretable in Q^- by a result of Vítěslav Švejdar. See [Šve07].

5. R is Top

In this section, we prove our main result.

Theorem 5.1. *For any locally finitely satisfiable RE theory T, we have $\mathsf{R} \rhd T$, via a one-dimensional, one-piece, parameter-free interpretation.*

The proof has the following structure. We construct a certain class of 'numbers' Good in R using a predicate α that describes the axiom set of T. Either Good provides a certain uniquely determined element $g^\star$ or not. If it does we can use $g^\star$ to uniquely specify an internally finite model z of a non-standardly finite part of T. Satisfaction in this model provides the desired interpretation of T. If $g^\star$ does not exist, this shows that Good is so rich that we can construct an interpretation of Q^- in Good. Hence, by Švejdar's result we have in interpretation of Q in Good. But as soon as we have Q, we have a whole range of well known techniques available to construct an interpretation of T. We will exhibit two such ways.

Suppose T is a locally finitely satisfiable RE theory. Let the axiom set of T be given by the Σ^0_1-formula α. We may assume that αy is of the form $\exists x\, \alpha_0 xy$, where α_0 is Δ_0. We can always arrange that $\alpha_0 mn$ implies $m < n$.

We need to speak about finite models and satisfaction in the context of number theory. Here we sketch one way how to do it. We code finite structures for the signature of T by numbers. We use the iterations of Cantor Pairing $(\cdot,\cdot)$ to implement finite sequences $(\cdot,\cdot,\ldots)$ of fixed standard length. We use Ackermann coding to represent finite sets. This means that we define $x \in y$ as $\exists u, v{\leq}y\ (y = u \cdot 2^{x+1} + v \wedge 2^x \leq v < 2^{x+1})$. (Remember that the graph of exponentiation is Δ_0-definable.) Suppose the signature of T is $\Sigma := \langle P, Q, \ldots\rangle$. We code a finite structure for Σ as a sequence $n := \langle m, p, q, \ldots\rangle$, where m stands for domain $\{0, \ldots, m-1\}$, and where p codes, in Ackermann style, a finite set of sequences with length the arity of P, etc. Just to make the argument run smoothly we stipulate that 0 is a code for the one-element model in which all atomic statements made with predicate symbols from the signature are false. We also assume that 0 is not a Gödel number of a formula.

We can find a Σ^0_1-formula $z, f \models a$ meaning: the model z and the assignment f satisfy the formula a. For any formula A we let A^y be the result of bounding all quantifiers in A by y. We define $\mathsf{comm}(u)$ as the conjunction of the following statements.

i. For all models $z < u$, the assignments for z on sets of variables w, with $w < u$, are closed under random reset for the domain of z. In other words, for a variable $w < u$, and $a < (z)_0$, $f[w : a]$ exists and, for all $w', b < u$, with w' a variable and $b < (z)_0$, we have:
$$f[w : a](w') = b \leftrightarrow ((w' \neq w \wedge fw' = b) \vee (w' = w \wedge b = a)).$$
ii. All formulas $a < u$, have a unique analysis into terms and formulas below u.
iii. We have the commutation conditions for formulas a and models z below u.

We find that, if $n \geq 2^{2^{c \cdot m^2}}$, for a sufficiently large standard c, then $\mathsf{comm}^n(m)$.

Let $\mathsf{good}(x)$ be a predicate coding the conjunction of (a) and (b).

a. There is a largest $y \leq x$, such that there is a minimal $z \leq x$, such that $\mathsf{comm}^x(\mathsf{max}(z, y))$, and, for any $a \leq y$, if $a \in \alpha^y$, then $z \models^x a$.
b. For all $u, v, w \leq x$, we have: $\mathsf{S}u \neq 0$, $\mathsf{S}u = \mathsf{S}v \rightarrow u = v$, $u = 0$ or $\exists q{\leq}x\ \mathsf{S}q = u$, $u + 0 = u$, $u + \mathsf{S}v = \mathsf{S}(u + v)$, $u \times 0 = 0$, $u \times \mathsf{S}v = u \times v + u$.

We work in $\mathsf{R}^\star$. Let Good be the virtual class of all g such that, for all $x \leq g$, $\mathsf{good}(x)$. Clearly, for every standard n, we can show that $\underline{n}$ is in Good. We note that Good is downwards closed w.r.t. $\leq$. Either Good is closed under successor or it is not.

We consider first the case that Good is not closed under successor. Consider g and g' such that $g, g' \in \mathsf{Good}$, $\mathsf{S}g \notin \mathsf{Good}$, $\mathsf{S}g' \notin \mathsf{Good}$. Suppose $g < g'$. Then by axiom R8, $\mathsf{S}g \leq g'$. So, by the downwards closure of Good, $\mathsf{S}g \in \mathsf{Good}$. Quod non. So, $g \not< g'$. Similarly, $g' \not< g$. Since, by R7, $\leq$ is linear, we find $g = g'$. So there is a unique element $g^\star$ in Good, such that $g^\star$ does not have a successor.

Since $g^\star$ is good, we can find a largest $y^\star \leq g^\star$, such that there is a minimal $z^\star \leq g^\star$, such that $\mathsf{comm}^{g^\star}(\mathsf{max}(z^\star, y^\star))$, and, for any $a \leq y^\star$, if $a \in \alpha^{y^\star}$, then $z^\star \models^{g^\star} a$.

Since T, is locally finite, for every standard n, we can show that $\underline{n} < y^\star$. It follows that for any standard axiom A of T, we have $z^\star \models^{g^\star} \ulcorner A \urcorner$.

To simplify inessentially, let's assume that T has just one binary predicate symbol P. We define a (parameter-free, one-dimensional, one-piece, identity-preserving) translation ν as follows:

- $\delta_\nu(v) :\leftrightarrow v < (z^\star)_0$,
- $P_\nu(v_0, v_1) :\leftrightarrow z^\star, f \models^{g^\star} \ulcorner P(v_0, v_1)\urcorner$, where f codes the assignment on just $\ulcorner v_0 \urcorner$ and $\ulcorner v_1 \urcorner$ such that $f\ulcorner v_0\urcorner = v_0$ and $f\ulcorner v_1 \urcorner = v_1$. This assignment exists since $\mathsf{comm}^{g^\star}(\mathsf{max}(z^\star, y^\star))$. We need only two resets to obtain f.

We can now prove, by external induction, that for any standard A, $z^\star \models^{g^\star} \ulcorner A \urcorner$ iff A^ν, since we have the commutation conditions for $\models^{g^\star}$ for formulas below $y^\star$. Hence, for any axiom A of T, we have A^ν.

Note that ν is parameter-free since the starred elements are definable. Also ν is identity preserving. Thus, we have shown that $\mathsf{R}^\star$ plus 'Good is not closed under successor' interprets T.

We turn to the case where Good is closed under successor. In other words, we work in $\mathsf{R}^\star$ plus 'Good is closed under successor'. Clearly 0 is in Good. Consider the following (parameter-free, one-dimensional, one-piece) translation γ for the language of Q^-.

- $\delta_\gamma(x) :\leftrightarrow \mathsf{Good}(x)$,
- $0_\gamma := 0$,
- $\mathsf{S}_\gamma x := \mathsf{S}x$,
- $\mathsf{A}_\gamma xyz :\leftrightarrow x, y, z \in \mathsf{Good} \wedge x + y = z$,
- $\mathsf{M}_\gamma xyz :\leftrightarrow x, y, z \in \mathsf{Good} \wedge x \times y = z$.

We evidently find $(\mathsf{Q}^-)^\gamma$. Thus we have shown that $\mathsf{R}^\star$ plus 'Good is closed under successor' interprets Q^-. Vítěslav Švejdar shows that Q^- interprets Q. See [Šve07]. So, we have $\mathsf{R}^\star$ plus 'Good is closed under successor' interprets Q. It follows that it is sufficient to show $\mathsf{Q} \rhd T$. If so, we have both $\mathsf{R}^\star$ plus 'Good is not closed under successor' interprets T, and $\mathsf{R}^\star$ plus 'Good is closed under successor' interprets T. Then, we can form a disjunctive interpretation to show that $\mathsf{R}^\star \rhd T$.

Finally, we prove that Q interprets T. There are two ways to do it. Here is the first. We know that Q interprets a convenient theory like Buss' S^1_2 on a definable cut. (See e.g. [HP91].) So it is sufficient to show that S^1_2 interprets T. As before, we can write down a satisfaction predicate for finite models. We can find definable cut J (downwards closed w.r.t. $\leq$, closed under S, $+$, $\times$ and ω_1, or equivalently $\#$.) such that S^1_2 proves that we have the commutation conditions for formulas in J. The crux here is that the witnesses for satisfaction need not be in J.[6] Let's write $\alpha \restriction y$ for the set of all axioms witnessed below y. Using this, we can find a definable cut $J^\star$, such that:

$$\mathsf{S}^1_2 \vdash \forall y, z{\in}J^\star\ (z \models \alpha \restriction y \to \forall p, a{\in}J^\star\ (\mathsf{proof}_{\alpha\restriction y}(p, a) \to z \models a))$$

From this it follows that, for every n, $\mathsf{S}^1_2 \vdash \mathsf{con}^{J^\star}(\alpha \restriction n)$. In more sloppy notation: $\mathsf{S}^1_2 \vdash \mathsf{con}^{J^\star}(T \restriction n)$. We can now use the Henkin-Feferman argument to build an interpretation of T in S^1_2. (See e.g. [Vis08].) In terms of that paper, we have shown: $Q \rhd \mho(T)$.) This interpretation is one-dimensional, one-piece and parameter free. Moreover, by [Vis09b], Corollary 6.1, we find that we can make the interpretation identity-preserving.

[6]In fact, the argument is closely analogous with the construction of a Σ^0_1-truth predicate that works for formulas in a cut J.

Remark 5.2. Since the axiomatization of R is reasonably simple, we can do more. One can show that $\mathsf{S}^1_2 \vdash \mathsf{con}^{J^\star}(R)$.

We can use this insight to produce an example of a theory U that is not locally finitely satisfiable but such that Q still interprets $\mathsf{Q} + \mathsf{con}(U)$. Reason in S^1_2. We either have $\mathsf{con}^{J^\star}(\mathsf{Q})$ or $\mathsf{incon}^{J^\star}(\mathsf{Q})$. In the first case, we find, by Pudlák's version of the second incompleteness theorem, that $\mathsf{con}^{J^\star}(\mathsf{Q} + \mathsf{incon}(\mathsf{Q}))$. So, a fortiori, we have $\mathsf{con}^{J^\star}(\mathsf{R} + \mathsf{incon}(\mathsf{Q}))$. In the second case, we have $\mathsf{incon}^{J^\star}(\mathsf{Q})$. We find, by $\exists\Sigma^{\mathsf{b}}_1$-completeness, that $\mathsf{con}^{J^\star}(\mathsf{R} + \mathsf{incon}(\mathsf{Q}))$. So in both cases we have $\mathsf{con}^{J^\star}(\mathsf{R} + \mathsf{incon}(\mathsf{Q}))$. We may conclude that Q interprets $\mathsf{Q} + \mathsf{con}(\mathsf{R} + \mathsf{incon}(\mathsf{Q}))$. Evidently, $\mathsf{R} + \mathsf{incon}(\mathsf{Q})$ is not locally finitely satisfiable. ❏

We turn to the second proof of the interpretability of T in Q. By a result of Wilkie, we know that Q interprets $\mathrm{I}\Delta_0 + \Omega_1$ on a definable cut. So, it is sufficient to show the interpretability of T in $\mathrm{I}\Delta_0 + \Omega_1$. Let J be a definable cut such that we have $\mathrm{I}\Delta_0 + \Omega_1 \vdash \forall x {\in} J\ 2^{2^{cx^2}}\downarrow$. As is well-known we have $(\mathrm{I}\Delta_0 + \Omega_1) \rhd (\mathrm{I}\Delta_0 + \Omega_1 + \mathsf{incon}^J(\mathsf{Q}))$. This interpretation uses the Henkin-Feferman construction and is one-dimensional, one-piece and parameter-free. We can arrange it to be also identity-preserving. So it is sufficient to show that $W := \mathrm{I}\Delta_0 + \Omega_1 + \mathsf{incon}^J(\mathsf{Q})$ interprets T.

We work in W. Let $p^\star$ be the smallest proof of $\mathsf{incon}(\mathsf{Q})$. Let $y^\star$ be the largest number below $p^\star$, such that there is a model z below $p^\star$ that satisfies all axioms of α witnessed below $y^\star$. The number $y^\star$ exists since we can bound the existential quantifier in the satisfaction predicate by $2^{2^{c(p^\star)^2}}$. Now we take the smallest such model $z^\star$ for the given $y^\star$. Again we use the above bounding argument to show that $z^\star$ exists. We use $z^\star$ to construct the desired interpretation of T in a way that is completely analogous to our earlier construction of γ.

Since all interpretations we used along the way are one-dimensional, one-piece, parameter-free and identity-preserving, their composition also has these desired properties.

References

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

[JS83] J. P. Jones and J.C. Shepherdson. Variants of Robinson's essentially undecidable theory R. *Archiv für Mathematische Logik und Grundlagenforschung*, 23:61–64, 1983.

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

[Nel86] E. Nelson. *Predicative arithmetic*. Princeton University Press, Princeton, 1986.

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

[Šve07] V. Švejdar. An interpretation of Robinson's Arithmetic in Grzegorczyk's weaker variant. *Fundamenta Informaticae*, 81:347–354, 2007.

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

[Vau62] R.L. Vaught. On a theorem of Cobham concerning undecidable theories. In E. Nagel, P. Suppes, and A. Tarski, editors, *Logic, Methodology and Philosophy of Science. Proceedings of the 1960 International Congress*, pages 14–25. Stanford University Press, Stanford, 1962.

[Vis08] A. Visser. Can we make the Second Incompleteness Theorem coordinate free. Logic Group Preprint Series 267, Department of Philosophy, Utrecht University, Heidelberglaan 8, 3584 CS Utrecht, ⟨http://www.phil.uu.nl/preprints/lgps/⟩, 2008.

[Vis09a] A. Visser. Cardinal arithmetic in the style of baron von Münchhausen. *Review of Symbolic Logic*, 2(3):??–??, 2009. To appear.

[Vis09b] A. Visser. The predicative Frege hierarchy. *Annals of Pure and Applied Logic*, 160(2):129–153, 2009. doi: 10.1016/j.apal.2009.02.001.

Department of Philosophy, Utrecht University, Heidelberglaan 8, 3584 CS Utrecht, The Netherlands

E-mail address: Albert.Visser@phil.uu.nl