The theory of hereditarily bounded sets

Emil Jeřábek*

Institute of Mathematics, Czech Academy of Sciences

Žitná 25, 115 67 Praha 1, Czech Republic, email: `jerabek@math.cas.cz`

October 4, 2021

Abstract

We show that for any $k \in \omega$, the structure $\langle H_k, \in\rangle$ of sets that are hereditarily of size at most k is decidable. We provide a transparent complete axiomatization of its theory, a quantifier elimination result, and tight bounds on its computational complexity. This stands in stark contrast to the structure $V_\omega = \bigcup_k H_k$ of hereditarily finite sets, which is well known to be bi-interpretable with the standard model of arithmetic $\langle \mathbb{N}, +, \cdot\rangle$.

1 Introduction

The *Vaught set theory VS*, originally introduced by Vaught [14], is a very rudimentary theory of sets: it is axiomatized by the schema

$$(\mathrm{V}_n) \qquad \forall x_0, \dots, x_{n-1}\, \exists y\, \forall t \Big(t \in y \leftrightarrow \bigvee_{i<n} t = x_i \Big)$$

for all $n \in \omega$, asserting that $\{x_i : i < n\}$ exists. It is one of the weakest known essentially undecidable theories; while Robinson's theory R, introduced in [12], is even weaker (in terms of interpretability), VS is appealing in the simplicity of its axioms, especially in the context of set theories where setting up an interpretation of an arithmetic theory such as R may be somewhat of a laborious task.

In contrast to full VS, the finite fragments VS_k (axiomatized by (V_0) and (V_k), which imply (V_m) for all $m \le k$) are *not* essentially undecidable, but the reason for this is a bit indirect: for each k, VS_k is interpretable in any theory with a pairing function, and it is known that there exist decidable consistent theories with pairing.

The first such theories were constructed by Malcev [9, 10]: he proved the decidability of theories of *locally free algebras*, which are essentially the first-order theories of term algebras in a given signature. His results also apply to free algebras with function symbols constrained to be symmetric w.r.t. prescribed groups of permutations of the arguments. As a special case, *acyclic*

*Supported by grant 19-05497S of GA ČR. The Institute of Mathematics of the Czech Academy of Sciences is supported by RVO: 67985840.

pairing functions are locally free algebras with a single binary function; e.g., the pairing function $2^x 3^y$ on $\mathbb{N}$ is acyclic, hence $\langle \mathbb{N}, 2^x 3^y \rangle$ is decidable. More generally, Tenney [13] proved that pairing functions that are acyclic up to a finite (or sufficiently well-behaved) set of exceptions have a decidable theory, including common pairing functions on $\mathbb{N}$ such as $2^x(2y+1)-1$, $\max\{x^2, y^2+x\}+y$, or Cantor's function $C(x,y) = \binom{x+y+1}{2} + x$. (The decidability of $\langle \mathbb{N}, C \rangle$ was reproved in [4] using Malcev's results.) Decidable structures with pairing may include more arithmetic functions: Cégielski and Richard observed in [5] that pairing functions such as $2^x + 2^{x+y}$ are definable in $\langle \mathbb{N}, +, 2^x \rangle$, which is decidable due to Semenov [11], and in a tour de force [6], they proved the decidability of $\langle \mathbb{N}, C, S \rangle$ (while other related structures, including $\langle \mathbb{N}, C, < \rangle$, $\langle \mathbb{N}, C, + \rangle$, and $\langle \mathbb{N}, C, \cdot \rangle$, are undecidable).

For more background on theories with "containers" such as pairs, sets, and sequences, see Visser [15].

While the results above confirm that finite fragments of the Vaught set theory are not essentially undecidable, the decidable extensions of VS_k we get from interpretation in theories of pairing are quite unnatural when we think of them as set theories: for example, they will contradict extensionality, which is arguably the most characteristic principle distinguishing sets from other kinds of objects. Thus, it might be worthwhile to see if we can find decidable extensions of VS_k that are easier to understand.

One of the simplest—and perhaps most natural—models of VS_k is the structure[1] $\mathbf{H}_k = \langle H_k, \in \rangle$ of *sets hereditarily of size at most* k; that is, H_k is the smallest family of sets such that every subset of H_k of cardinality $\le k$ is a member of H_k:

$$\forall x \left(x \subseteq H_k \land |x| \le k \implies x \in H_k \right).$$

Equivalently, H_k consists of (well-founded) sets x such that x itself, and all elements of its transitive closure, have cardinality $\le k$. The better known family of *hereditarily finite* sets V_ω includes each H_k, and in fact, $V_\omega = \bigcup_{k \in \omega} H_k$. We observe that $\mathbf{H}_k$ is a *minimal* model of VS_k, in that it embeds (as a transitive submodel) into any other model of VS_k; thus, $\mathbf{H}_k$ is canonically associated with VS_k.

The main purpose of this paper is to show that $\mathrm{Th}(\mathbf{H}_k)$ is decidable, providing an explicit natural example of a decidable extension of VS_k. We present a transparent recursive axiomatization of $\mathrm{Th}(\mathbf{H}_k)$, and a characterization of elementary equivalence of tuples in models of $\mathrm{Th}(\mathbf{H}_k)$ in terms of isomorphism of transitive closures. Apart from the decidability of $\mathbf{H}_k$, this yields a quantifier elimination result (every formula is equivalent to a Boolean combination of bounded existential formulas). We also establish that $\mathrm{Th}(\mathbf{H}_k)$ is stable, and it is not finitely axiomatizable. In Section 3, we investigate in more detail the computational complexity of $\mathrm{Th}(\mathbf{H}_k)$: we give an algorithm deciding $\mathrm{Th}(\mathbf{H}_k)$ whose running time closely matches a general lower bound on the complexity of theories with pairing by Ferrante and Rackoff [7], and its variant that has much lower complexity for sentences with a small number of quantifier alternations.

The properties of $\mathbf{H}_k$ may be contrasted with the structure $\langle V_\omega, \in \rangle$, which is bi-interpretable

[1]The standard notation in set theory is that, for a (usually regular infinite) cardinal κ, H_κ consists of sets hereditarily of cardinality $<\kappa$, thus our H_k would be denoted H_{k+1}. We decided to violate this convention as it seems to be more confusing than helpful in the finite case.

with $\langle \mathbb{N}, +, \cdot \rangle$ by Ackermann [1], and as such it is heavily undecidable, and its quantifier alternation hierarchy is proper.

Let us remark that while we formulate most results so that they apply to all $k \in \omega$, the cases $k = 0, 1$ are somewhat degenerate: $\mathbf{H}_0$ is a one-element structure, and $\mathbf{H}_1$ is definitionally equivalent to $\langle H_1, \varnothing, \{-\} \rangle \simeq \langle \mathbb{N}, 0, S \rangle$. Moreover, the case $k = 2$ can be reduced to Malcev's results: $\mathbf{H}_2$ is definitionally equivalent to the structure $\langle H_2, \varnothing, \{-, -\} \rangle$, which is a free algebra with a constant and a commutative binary operation. A similar reduction does not seem possible for $k \geq 3$, as the set builder operation $\{x_0, \ldots, x_{k-1}\}$ has peculiar symmetries such as $\{x, x, y\} = \{x, y, y\}$ that cannot be expressed by mere permutations of arguments.

2 Completeness and decidability

The main result of this section is the decidability of $\mathbf{H}_k$. Our strategy is to propose a recursively axiomatized theory S_k, true in $\mathbf{H}_k$, and prove its completeness: this implies that S_k is decidable and $S_k = \mathrm{Th}(\mathbf{H}_k)$. Without further ado, here is the definition of S_k.

Definition 2.1 Let $k \in \omega$. The theory S_k in the language of set theory $\langle \in \rangle$ is axiomatized by (V_0), (V_k), the extensionality axiom

$$\text{(E)} \qquad \forall x, y \,\bigl(\forall t\, (t \in x \leftrightarrow t \in y) \to x = y\bigr),$$

the boundedness axiom

$$\text{(B}_k\text{)} \qquad \forall x, u_0, \ldots, u_k \,\Bigl(\bigwedge_{i \leq k} u_i \in x \to \bigvee_{i < j \leq k} u_i = u_j\Bigr)$$

postulating that all sets have at most k elements, and the axioms

$$\text{(C}_n\text{)} \qquad \forall x_0, \ldots, x_n \,\neg\Bigl(\bigwedge_{i < n} x_i \in x_{i+1} \wedge x_n = x_0\Bigr)$$

for all $n \in \omega$, $n \geq 1$, prohibiting finite $\in$-cycles.

Clearly, $\mathbf{H}_k \vDash S_k$. We aim to show that S_k is complete; we will prove this by an Ehrenfeucht–Fraïssé argument, which will more generally provide a characterization of elementary equivalence of finite tuples in models of S_k. Let us first agree on basic notation concerning models.

Definition 2.2 As a general notational convention, we will denote first-order structures by bold-face letters (possibly decorated). The domain of a structure will be denoted by the same letter, but in italics, and the basic relations and functions of a structure carry the name of the structure as a superscript (this convention will also extend on a case-by-case basis to various defined concepts). For example, a typical model of the language of set theory will be denoted $\mathbf{A}$, in which case $\mathbf{A} = \langle A, \in^{\mathbf{A}} \rangle$.

We denote finite tuples (sequences) by letters with bars such as $\overline{a}$; then $\mathrm{lh}(\overline{a})$ denotes the length of $\overline{a}$, and the individual elements of $\overline{a}$ are a_i with $0 \leq i < \mathrm{lh}(\overline{a})$.

Let $\mathbf{A}$ and $\mathbf{B}$ be structures for the same language, and $\overline{a} \in A$, $\overline{b} \in B$ finite tuples of the same length $l = \mathrm{lh}(\overline{a}) = \mathrm{lh}(\overline{b})$. We write $\mathbf{A}, \overline{a} \equiv \mathbf{B}, \overline{b}$ if $\overline{a}$ and $\overline{b}$ satisfy the same formulas, and $\mathbf{A}, \overline{a} \equiv_n \mathbf{B}, \overline{b}$ if they satisfy the same formulas of quantifier rank at most n. We recall that the quantifier rank of a formula φ is defined inductively by

$$\begin{aligned} \mathrm{rk}(\varphi) &= 0, && \varphi \text{ quantifier-free}, \\ \mathrm{rk}\big(c(\varphi_0, \dots, \varphi_{k-1})\big) &= \max\big\{\mathrm{rk}(\varphi_i) : i < k\big\}, && c \in \{\wedge, \vee, \to, \neg\}, \\ \mathrm{rk}(Qx\, \varphi) &= \mathrm{rk}(\varphi) + 1, && Q \in \{\exists, \forall\}. \end{aligned}$$

If $f \colon A \to B$ and $X \subseteq A$, then $f[X]$ denotes the image $\{f(x) : x \in X\}$, and $f \restriction X \colon X \to B$ the restriction of f to X. If $\overline{a} \in A$ with $l = \mathrm{lh}(\overline{a})$, then $f(\overline{a})$ denotes the l-tuple $\overline{b}$ such that $b_i = f(a_i)$ for each $i < l$.

We also fix some notation and terminology specific to models of S_k. In particular, we intend to characterize the elementary equivalence relations $\equiv_n$ in terms of isomorphism of levels of transitive closures, hence we need to define the latter.

Definition 2.3 *Bounded quantifiers* in the language of set theory are introduced as the abbreviations

$$\begin{aligned} \exists y \in x\, \varphi &\equiv \exists y\, (y \in x \wedge \varphi), \\ \forall y \in x\, \varphi &\equiv \forall y\, (y \in x \to \varphi), \end{aligned}$$

where x and y are distinct variables. A formula is *bounded* if it is built from atomic formulas using Boolean connectives and bounded quantifiers.

If $\mathbf{A} \vDash S_k$ and $u \in A$, then

$$u^{\mathbf{A}} = \{v \in A : v \in^{\mathbf{A}} u\}$$

denotes the *extension* of u in $\mathbf{A}$. Conversely, if $r \leq k$ and $\{u_i : i < r\} \subseteq A$, then $\{u_i : i < r\}^{\mathbf{A}}$ or $\{u_0, \dots, u_{r-1}\}^{\mathbf{A}}$ denotes the $v \in A$ such that $v^{\mathbf{A}} = \{u_i : i < r\}$, which exists by ($\mathrm{V}_0$) or ($\mathrm{V}_k$), and is unique by (E). In particular, $\varnothing^{\mathbf{A}} = \{\}^{\mathbf{A}}$.

If $\overline{a} \in A$ and $l = \mathrm{lh}(\overline{a})$, we define levels of the *transitive closure* of $\overline{a}$ (as subsets of A) by

$$\begin{aligned} \mathrm{tc}_0^{\mathbf{A}}(\overline{a}) &= \{a_i : i < l\}, \\ \mathrm{tc}_{n+1}^{\mathbf{A}}(\overline{a}) &= \mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup \bigcup_{u \in \mathrm{tc}_n^{\mathbf{A}}(\overline{a})} u^{\mathbf{A}}, \\ \mathrm{tc}^{\mathbf{A}}(\overline{a}) &= \bigcup_{n \in \omega} \mathrm{tc}_n^{\mathbf{A}}(\overline{a}). \end{aligned}$$

We denote by $\mathbf{tc}_n^{\mathbf{A}}(\overline{a})$ the (possibly empty) structure $\langle \mathrm{tc}_n^{\mathbf{A}}(\overline{a}), \in^{\mathbf{A}}, \overline{a} \rangle$, and likewise, $\mathbf{tc}^{\mathbf{A}}(\overline{a}) = \langle \mathrm{tc}^{\mathbf{A}}(\overline{a}), \in^{\mathbf{A}}, \overline{a} \rangle$.

Notice that $\mathrm{tc}_n^{\mathbf{A}}(\overline{a}) = \bigcup_{i<l} \mathrm{tc}_n^{\mathbf{A}}(a_i)$, and $\mathrm{tc}_n^{\mathbf{A}}(\overline{a})$ is finite: $|\mathrm{tc}_n^{\mathbf{A}}(\overline{a})| \leq l\, k^{\leq n}$, where

$$k^{\leq n} = \sum_{i=0}^{n} k^i = \begin{cases} \dfrac{k^{n+1} - 1}{k - 1}, & k \neq 1, \\ n + 1, & k = 1. \end{cases}$$

Also, for any fixed n and l, there is a formula $\varphi(\overline{x}, y)$ with $\mathrm{lh}(\overline{x}) = l$ that defines the relation $y \in \mathrm{tc}_n^{\mathbf{A}}(\overline{x})$ in every model $\mathbf{A} \vDash S_k$. Finally, we define

$$\begin{aligned}\mathbf{A}, \overline{a} \sim_n \mathbf{B}, \overline{b} &\iff \mathbf{tc}_n^{\mathbf{A}}(\overline{a}) \simeq \mathbf{tc}_n^{B}(\overline{b}),\\ \mathbf{A}, \overline{a} \sim \ \mathbf{B}, \overline{b} &\iff \mathbf{tc}^{A}(\overline{a}) \simeq \mathbf{tc}^{B}(\overline{b}).\end{aligned}$$

We first observe basic properties of morphisms on transitive closures.

Lemma 2.4 *Let* $\mathbf{A}, \mathbf{B} \vDash S_k$, $\overline{a} \in A$, $\overline{b} \in B$, $\mathrm{lh}(\overline{a}) = \mathrm{lh}(\overline{b})$, *and* $n > 0$.

(i) *If* $f\colon \mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \to B$ *is a mapping such that* $f(\overline{a}) = \overline{b}$ *and*

$$\forall u \in \mathrm{tc}_{n-1}^{\mathbf{A}}(\overline{a})\, f(u) = \{f(t) : t \in^{\mathbf{A}} u\}^{\mathbf{B}}, \tag{1}$$

then $f[\mathrm{tc}_m^{\mathbf{A}}(\overline{a})] = \mathrm{tc}_m^{\mathbf{B}}(\overline{b})$ *for all* $m \le n$.

(ii) *Any* $f\colon \mathbf{tc}_n^{\mathbf{A}}(\overline{a}) \simeq \mathbf{tc}_n^{\mathbf{B}}(\overline{b})$ *satisfies* (1), *thus* $f \restriction \mathrm{tc}_m^{\mathbf{A}}(\overline{a})\colon \mathbf{tc}_m^{\mathbf{A}}(\overline{a}) \simeq \mathbf{tc}_m^{\mathbf{B}}(\overline{b})$ *for all* $m \le n$.

Proof:

(i): By induction on m. The case $m = 0$ holds. Assume $f[\mathrm{tc}_m^{\mathbf{A}}(\overline{a})] = \mathrm{tc}_m^{\mathbf{B}}(\overline{b})$ and $m < n$. If $t \in \mathrm{tc}_{m+1}^{\mathbf{A}}(\overline{a})$, then $t \in^{\mathbf{A}} u$ for some $u \in \mathrm{tc}_m^{\mathbf{A}}(\overline{a})$, thus $f(t) \in^{\mathbf{B}} f(u) \in \mathrm{tc}_m^{\mathbf{B}}(\overline{b})$ by (1) and the induction hypothesis, which means $f(t) \in \mathrm{tc}_{m+1}^{\mathbf{B}}(\overline{b})$. Conversely, if $s \in \mathrm{tc}_{m+1}^{\mathbf{B}}(\overline{b})$, we have $s \in^{\mathbf{B}} v$ for some $v \in \mathrm{tc}_m^{\mathbf{B}}(\overline{b})$. By the induction hypothesis, there is $u \in \mathrm{tc}_m^{\mathbf{A}}(\overline{a})$ such that $f(u) = v$, thus $s = f(t)$ for some $t \in^{\mathbf{A}} u$ by (1), whence $t \in \mathrm{tc}_{m+1}^{\mathbf{A}}(\overline{a})$.

(ii): Let $u \in \mathrm{tc}_{n-1}^{\mathbf{A}}(\overline{a})$. We can prove $f(u) \in \mathrm{tc}_{n-1}^{\mathbf{B}}(\overline{b})$ as in (i). On the one hand, if $t \in^{\mathbf{A}} u$, then $t \in \mathrm{dom}(f)$, hence $f(t) \in^{\mathbf{B}} f(u)$. Now, on the other hand, if $s \in^{\mathbf{B}} f(u)$, then $s \in \mathrm{tc}_n^{\mathbf{B}}(\overline{b})$, which means that $s = f(t)$ for some $t \in \mathrm{tc}_n^{\mathbf{A}}(\overline{a})$. Then $f(t) \in^{\mathbf{B}} f(u)$ implies $t \in^{\mathbf{A}} u$. ☐

By definition, $\equiv\, = \bigcap_n \equiv_n$. It may not be a priori obvious that the same holds for the $\sim$ relation (which we aim to eventually prove to coincide with $\equiv$): e.g., the corresponding property fails for general pointed directed acyclic graphs. However, here it is true because axiom (B_k) ensures that the graphs are image-finite:

Lemma 2.5 *Let* $\mathbf{A}, \mathbf{B} \vDash S_k$, $\overline{a} \in A$, $\overline{b} \in B$, *and* $\mathrm{lh}(\overline{a}) = \mathrm{lh}(\overline{b})$. *Then* $\mathbf{A}, \overline{a} \sim \mathbf{B}, \overline{b}$ *if and only if* $\forall n \in \omega\, \mathbf{A}, \overline{a} \sim_n \mathbf{B}, \overline{b}$.

Proof: The left-to-right implication is clear. For the converse, Lemma 2.4 shows that the set T of all isomorphisms $f\colon \mathbf{tc}_n^{\mathbf{A}}(\overline{a}) \simeq \mathbf{tc}_n^{\mathbf{B}}(\overline{b})$, $n \in \omega$, forms a tree when ordered by inclusion, and the finiteness of tc_n implies that T is finitely branching. As such, T has an infinite branch by Kőnig's lemma; the union of the branch is then an isomorphism of $\mathbf{tc}^{\mathbf{A}}(\overline{a})$ to $\mathbf{tc}^{\mathbf{B}}(\overline{b})$. ☐

It is relatively straightforward to prove that $\mathbf{A}, \overline{a} \equiv \mathbf{B}, \overline{b}$ implies $\mathbf{A}, \overline{a} \sim \mathbf{B}, \overline{b}$: in view of the previous lemma, we only need to establish that the isomorphism types of the finite structures $\mathbf{tc}_n^{\mathbf{A}}(\overline{a})$ are definable. We do this below, including explicit bounds on the complexity of the defining formulas.

Lemma 2.6 *Let $\mathbf{A} \vDash S_k$, $\overline{a} \in A$, $l = \mathrm{lh}(\overline{a})$, and $n \in \omega$. Then there is a formula $\varphi_{\overline{a},n}(\overline{x})$ such that for any $\mathbf{B} \vDash S_k$ and any l-tuple $\overline{b} \in B$, we have*

$$\mathbf{B} \vDash \varphi_{\overline{a},n}(\overline{b}) \iff \mathbf{A}, \overline{a} \sim_n \mathbf{B}, \overline{b}.$$

Moreover, we may take $\varphi_{\overline{a},n}$ in the form $\psi(\overline{x}) \wedge \neg \bigvee_{i<m} \psi_i(\overline{x})$, where ψ and ψ_i are bounded existential formulas using at most $l(k^{\leq n} - 1)$ quantifiers each.

Proof: Let $\{a_i : l \leq i < r\}$ be an enumeration of $\mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \smallsetminus \{a_i : i < l\}$, where $r \leq l\, k^{\leq n}$, and for every $i \geq l$, there is $p(i) < i$ such that $a_i \in^{\mathbf{A}} a_{p(i)}$ (this can be arranged by enumerating elements of $\mathrm{tc}_{n'}^{\mathbf{A}}(\overline{a})$ before elements of $\mathrm{tc}_{n'+1}^{\mathbf{A}}(\overline{a}) \smallsetminus \mathrm{tc}_{n'}^{\mathbf{A}}(\overline{a})$, for each $n' < n$). Let θ be (the conjunction of) the diagram of $\{a_i : i < r\}$ with the structure induced from $\mathbf{A}$, and put

$$\psi(x_0, \ldots, x_{l-1}) = \exists x_l \in x_{p(l)}\, \exists x_{l+1} \in x_{p(l+1)}\, \ldots\, \exists x_{r-1} \in x_{p(r-1)}\, \theta(x_0, \ldots, x_{r-1}).$$

Then for any $\mathbf{B} \vDash S_k$ and $\overline{b} \in B$,

$$\mathbf{B} \vDash \psi(\overline{b}) \iff \mathbf{tc}_n^{\mathbf{A}}(\overline{a}) \mathrel{\widetilde{\subseteq}} \mathbf{tc}_n^{\mathbf{B}}(\overline{b}),$$

where $\mathbf{M} \mathrel{\widetilde{\subseteq}} \mathbf{N}$ denotes that there exists an embedding $f \colon \mathbf{M} \to \mathbf{N}$. Let $\{\mathbf{M}_i : i < m\}$ be an enumeration (up to isomorphism) of all structures of the form $\mathbf{tc}_n^{\mathbf{C}}(\overline{c})$ that do not embed into $\mathbf{tc}_n^{\mathbf{A}}(\overline{a})$, and as above, let ψ_i be a bounded existential formula in at most $l\, k^{\leq n}$ variables such that

$$\mathbf{B} \vDash \psi_i(\overline{b}) \iff \mathbf{M}_i \mathrel{\widetilde{\subseteq}} \mathbf{tc}_n^{\mathbf{B}}(\overline{b}).$$

Then $\varphi_{\overline{a},n} = \psi \wedge \neg \bigvee_{i<m} \psi_i$ satisfies

$$\begin{aligned}
\mathbf{B} \vDash \varphi_{\overline{a},n}(\overline{b}) &\iff \mathbf{tc}_n^{\mathbf{A}}(\overline{a}) \mathrel{\widetilde{\subseteq}} \mathbf{tc}_n^{\mathbf{B}}(\overline{b}) \wedge \forall i < m\, \mathbf{M}_i \mathrel{\widetilde{\nsubseteq}} \mathbf{tc}_n^{\mathbf{B}}(\overline{b}) \\
&\iff \mathbf{tc}_n^{\mathbf{A}}(\overline{a}) \mathrel{\widetilde{\subseteq}} \mathbf{tc}_n^{\mathbf{B}}(\overline{b}) \wedge \mathbf{tc}_n^{\mathbf{B}}(\overline{b}) \mathrel{\widetilde{\subseteq}} \mathbf{tc}_n^{\mathbf{A}}(\overline{a}) \\
&\iff \mathbf{tc}_n^{\mathbf{A}}(\overline{a}) \simeq \mathbf{tc}_n^{\mathbf{B}}(\overline{b}),
\end{aligned}$$

using the fact that if $\mathbf{M}$ and $\mathbf{N}$ are finite structures such that $\mathbf{M} \mathrel{\widetilde{\subseteq}} \mathbf{N}$ and $\mathbf{N} \mathrel{\widetilde{\subseteq}} \mathbf{M}$, then $\mathbf{M} \simeq \mathbf{N}$. □

Corollary 2.7 *Let $\mathbf{A}, \mathbf{B} \vDash S_k$, $\overline{a} \in A$, $\overline{b} \in B$, and $l = \mathrm{lh}(\overline{a}) = \mathrm{lh}(\overline{b})$. Then $\mathbf{A}, \overline{a} \equiv \mathbf{B}, \overline{b}$ implies $\mathbf{A}, \overline{a} \sim \mathbf{B}, \overline{b}$. More precisely, $\mathbf{A}, \overline{a} \equiv_{l(k^{\leq n}-1)} \mathbf{B}, \overline{b}$ implies $\mathbf{A}, \overline{a} \sim_n \mathbf{B}, \overline{b}$.* □

It is more difficult to show the converse implication $\mathbf{A}, \overline{a} \sim \mathbf{B}, \overline{b} \implies \mathbf{A}, \overline{a} \equiv \mathbf{B}, \overline{b}$. We will do it by an Ehrenfeucht–Fraïssé argument: that is, we will prove that if $\mathbf{A}, \overline{a} \sim_m \mathbf{B}, \overline{b}$ for m sufficiently larger than n, then any extension of $\overline{a}$ to $\overline{a}, c$ can be matched by an extension of $\overline{b}$ to $\overline{b}, d$ so that $\mathbf{A}, \overline{a}, c \sim_n \mathbf{B}, \overline{b}, d$. This is the content of the crucial Lemma 2.9 below. However, we start with a little technical result that will be needed in its proof.

Lemma 2.8 *Let $\mathbf{A} \vDash S_k$, $\overline{a} \in A$, and $n, r \in \omega$, where $k \geq 1$. There exists $\{v_i : i < r\} \subseteq A$ such that*

- $v_i \notin \mathrm{tc}^{\mathbf{A}}(\overline{a})$,

- $i \neq j \implies v_i \notin \mathrm{tc}_n^{\mathbf{A}}(v_j)$,

for all $i, j < r$.

Proof: We may assume that $l = \mathrm{lh}(\overline{a}) > 0$. The acyclicity of $\in^{\mathbf{A}}$ implies that the relation $x \in \mathrm{tc}^{\mathbf{A}}(y)$ (which is the reflexive transitive closure of $\in^{\mathbf{A}}$) is a partial order, hence its restriction to any nonempty finite set has a maximal element. That is, we can find $a \in \{a_i : i < l\}$ such that $a \notin \mathrm{tc}^{\mathbf{A}}(a_i)$ for any $a_i \neq a$. Then $\{a\}^{\mathbf{A}} \notin \mathrm{tc}^{\mathbf{A}}(\overline{a})$, which implies that $v_i = \{a\}^{1+(n+1)i}$ have the required properties, where $\{a\}^0 = a$, $\{a\}^{t+1} = \{\{a\}^t\}^{\mathbf{A}}$. (If $k \geq 2$, we may even ensure the stronger condition $v_i \notin \mathrm{tc}^{\mathbf{A}}(v_j)$ for $j \neq i$, by putting $v_i = \{\{a\}^{i+1}, \{a\}^{i+2}\}^{\mathbf{A}}$.) $\Box$

Lemma 2.9 *Let* $\mathbf{A}, \mathbf{B} \vDash S_k$, $\overline{a} \in A$, $\overline{b} \in B$, $l = \mathrm{lh}(\overline{a}) = \mathrm{lh}(\overline{b})$, *and* $n > 0$. *If* $\mathbf{A}, \overline{a} \sim_{k^{\leq n}+n} \mathbf{B}, \overline{b}$, *then for every* $c \in A$, *there exists* $d \in B$ *such that* $\mathbf{A}, \overline{a}, c \sim_{n-1} \mathbf{B}, \overline{b}, d$.

Proof: If $k = 0$, the conclusion of the lemma holds trivially as $|A| = |B| = 1$, hence we may assume $k \geq 1$. Put $N = k^{\leq n} + n$, and fix $f\colon \mathbf{tc}_N^{\mathbf{A}}(\overline{a}) \simeq \mathbf{tc}_N^{\mathbf{B}}(\overline{b})$. Let C be the smallest subset of $\mathrm{tc}_n^{\mathbf{A}}(c) \smallsetminus \mathrm{tc}_n^{\mathbf{A}}(\overline{a})$ satisfying the inductive condition

$$u^{\mathbf{A}} \subseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup C \implies u \in C$$

for $u \in \mathrm{tc}_n^{\mathbf{A}}(c) \smallsetminus \mathrm{tc}_n^{\mathbf{A}}(\overline{a})$. We can extend $f \restriction \mathrm{tc}_n^{\mathbf{A}}(\overline{a})$ uniquely to a mapping $g\colon \mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup C \to B$ such that

$$g(u) = \{g(t) : t \in^{\mathbf{A}} u\}^{\mathbf{B}}$$

for all $u \in C$. Let $\{u_i : i < r\}$ be an injective enumeration of

$$\{u \in \mathrm{tc}_n^{\mathbf{A}}(c) \smallsetminus \mathrm{tc}_n^{\mathbf{A}}(\overline{a}) : u^{\mathbf{A}} \nsubseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a}, c)\}.$$

Using Lemma 2.8, we can find $\{v_i : i < r\} \subseteq B$ such that

(i) $v_i \notin \mathrm{tc}_N^{\mathbf{B}}(\overline{b}) \cup g[C]$,

(ii) $i \neq j \implies v_i \notin \mathrm{tc}_N^{\mathbf{B}}(v_j)$,

for all $i, j < r$. Since $\in^{\mathbf{A}}$ is acyclic, and therefore well-founded on the finite set $\mathrm{tc}_n^{\mathbf{A}}(c)$, we can construct using well-founded recursion a unique mapping $g\colon \mathrm{tc}_n^{\mathbf{A}}(\overline{a}, c) \to B$ such that

$$g(u) = \begin{cases} f(u), & u \in \mathrm{tc}_n^{\mathbf{A}}(\overline{a}), \\ v_i, & u = u_i, \\ \{g(t) : t \in^{\mathbf{A}} u\}^{\mathbf{B}}, & u \in \mathrm{tc}_n^{\mathbf{A}}(c) \smallsetminus \mathrm{tc}_n^{\mathbf{A}}(\overline{a}), u^{\mathbf{A}} \subseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a}, c). \end{cases}$$

(This agrees with the original definition of g on $\mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup C$, hence keeping the same name will not lead to confusion. The reason for this slightly awkward two-stage construction of g is that we could not define the whole g right away as it depends on the choice of $\{v_i : i < r\}$, which in turn depends on $g \restriction C$.) Using Lemma 2.4 and the definition of g, the condition

$$g(u) = \{g(t) : t \in^{\mathbf{A}} u\}^{\mathbf{B}} \tag{2}$$

holds for all $u \in \mathrm{tc}_n^{\mathbf{A}}(\overline{a},c)$ such that $u^{\mathbf{A}} \subseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a},c)$. In particular, it holds for all $u \in \mathrm{tc}_{n-1}^{\mathbf{A}}(\overline{a},c)$, hence Lemma 2.4 implies $g[\mathrm{tc}_{n-1}^{\mathbf{A}}(\overline{a},c)] = \mathrm{tc}_{n-1}^{\mathbf{B}}(\overline{b},d)$, where $d = g(c)$.

We claim that g is injective. Assuming for the moment that this is true, let us show that $g \restriction \mathrm{tc}_{n-1}^{\mathbf{A}}(\overline{a},c)\colon\ \mathbf{tc}_{n-1}^{\mathbf{A}}(\overline{a},c) \simeq \mathbf{tc}_{n-1}^{\mathbf{B}}(\overline{b},d)$. If $t,u \in \mathrm{tc}_{n-1}^{\mathbf{A}}(\overline{a},c)$, then u satisfies (2). Thus, on the one hand, $t \in^{\mathbf{A}} u$ implies $g(t) \in^{\mathbf{B}} g(u)$; on the other hand, if $g(t) \in^{\mathbf{B}} g(u)$, then $g(t) = g(t')$ for some $t' \in^{\mathbf{A}} u$, and we have $t = t'$ by injectivity, hence $t \in^{\mathbf{A}} u$.

It remains to prove the injectivity of g. Assume for contradiction that there are $x, y \in \mathrm{tc}_n^{\mathbf{A}}(\overline{a},c)$ such that $x \neq y$, but $g(x) = g(y)$. Since $\in^{\mathbf{A}}$ is well-founded on $\mathrm{tc}_n^{\mathbf{A}}(\overline{a},c)$, we may take x to be $\in^{\mathbf{A}}$-minimal for which such a y exists.

If $x^{\mathbf{A}}, y^{\mathbf{A}} \subseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a},c)$ so that (2) holds for both x and y, there is $x' \in^{\mathbf{A}} x$ such that $x' \notin^{\mathbf{A}} y$, or $y' \in^{\mathbf{A}} y$ such that $y' \notin^{\mathbf{A}} x$. In the former case, $g(x') \in^{\mathbf{B}} g(x) = g(y)$, hence $g(x') = g(y')$ for some $y' \in^{\mathbf{A}} y$, and necessarily $x' \neq y'$; this contradicts the minimality of x. The other case is symmetric.

Thus, $x^{\mathbf{A}} \nsubseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a},c)$ or $y^{\mathbf{A}} \nsubseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a},c)$. By swapping x and y if necessary (dropping the minimality assumption, which is no longer needed), we may assume the latter. We distinguish two cases.

Case 1: $y = u_i$ for some $i < r$. Thus, $g(x) = v_i$ and $x \neq u_i$. We cannot have $x \in \mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup C$ because of (i), hence $x \in \mathrm{tc}_n^{\mathbf{A}}(c) \smallsetminus (\mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup C)$. Put $x_0 = x$. Either $x_0 = u_j$ for some j, or $x_0^{\mathbf{A}} \subseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a},c)$, while $x_0^{\mathbf{A}} \nsubseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup C$; thus, there is $x_1 \in^{\mathbf{A}} x_0$ such that $x_1 \in \mathrm{tc}_n^{\mathbf{A}}(c) \smallsetminus (\mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup C)$, and we can continue in the same way. By acyclicity of $\in^{\mathbf{A}}$, the process has to stop after less than $|\mathrm{tc}_n^{\mathbf{A}}(c)| \leq k^{\leq n}$ steps; that is, we can construct a sequence $x_0, \dots, x_s \in \mathrm{tc}_n^{\mathbf{A}}(c) \smallsetminus (\mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup C)$ such that $s < k^{\leq n}$, $x_s \in^{\mathbf{A}} \cdots \in^{\mathbf{A}} x_1 \in^{\mathbf{A}} x_0$, $x_i^{\mathbf{A}} \subseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a},c)$ for each $i < s$, and $x_s^{\mathbf{A}} \nsubseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a},c)$, which means $x_s = u_j$ for some $j < r$. But then

$$v_j = g(x_s) \in^{\mathbf{B}} \cdots \in^{\mathbf{B}} g(x_1) \in^{\mathbf{B}} g(x_0) = v_i$$

by (2), i.e., $v_j \in \mathrm{tc}_s^{\mathbf{B}}(v_i)$. By condition (ii), this is only possible if $j = i$, and then $s = 0$ by acyclicity of $\in^{\mathbf{B}}$. Thus, $x = u_i$ after all, a contradiction.

Case 2: $y \in \mathrm{tc}_n^{\mathbf{A}}(\overline{a})$. We cannot have $x \in \mathrm{tc}_n^{\mathbf{A}}(\overline{a})$ as f is injective. If $x \notin \mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup C$, then the argument in Case 1 shows that $v_j \in \mathrm{tc}_{k^{\leq n}}^{\mathbf{B}}(g(x))$ for some $j < r$, while $g(x) = f(y) \in \mathrm{tc}_n^{\mathbf{B}}(\overline{b})$, thus $v_j \in \mathrm{tc}_N^{\mathbf{B}}(\overline{b})$, contradicting (i). The only remaining possibility is $x \in C$. Put $x_0 = x$ and $y_0 = y$. We have $x_0^{\mathbf{A}} \subseteq \mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup C$, thus x_0 satisfies (2), while

$$f(y_0) = \{f(t) : t \in^{\mathbf{A}} y_0\}^{\mathbf{B}}$$

by Lemma 2.4. Thus, the same argument as above shows that there are $x_1 \in^{\mathbf{A}} x_0$ (whence $x_1 \in \mathrm{tc}_n^{\mathbf{A}}(\overline{a}) \cup C$) and $y_1 \in^{\mathbf{A}} y_0$ (whence $y_1 \in \mathrm{tc}_{n+1}^{\mathbf{A}}(\overline{a})$) such that $g(x_1) = f(y_1)$ and $x_1 \neq y_1$. If $x_1 \in C$, we may continue in the same way, but the acyclicity of $\in^{\mathbf{A}}$ again implies that the process has to stop: that is, we construct sequences $x_0, \dots, x_s$ and $y_0, \dots, y_s$ such that $s \leq |C| \leq k^{\leq n}$, $x_s \in^{\mathbf{A}} \cdots \in^{\mathbf{A}} x_1 \in^{\mathbf{A}} x_0$, $x_i \in C$ for each $i < s$, $x_s \in \mathrm{tc}_n^{\mathbf{A}}(\overline{a})$, $y_s \in^{\mathbf{A}} \cdots \in^{\mathbf{A}} y_1 \in^{\mathbf{A}} y_0$ (thus $y_i \in \mathrm{tc}_{n+i}^{\mathbf{A}}(\overline{a}) \subseteq \mathrm{tc}_N^{\mathbf{A}}(\overline{a})$), $x_i \neq y_i$ for each $i \leq s$, and $g(x_i) = f(y_i)$. But then $f(x_s) = f(y_s)$ contradicts the injectivity of f. This completes the proof. $\square$

We can now put everything together to obtain the desired characterization of elementary equivalence.

Theorem 2.10 *Let* $\mathbf{A}, \mathbf{B} \vDash S_k$, $\overline{a} \in A$, $\overline{b} \in B$, *and* $l = \mathrm{lh}(\overline{a}) = \mathrm{lh}(\overline{b})$. *Then*

$$\mathbf{A}, \overline{a} \equiv \mathbf{B}, \overline{b} \iff \mathbf{A}, \overline{a} \sim \mathbf{B}, \overline{b}.$$

More precisely, for all $n \in \omega$,

$$\mathbf{A}, \overline{a} \equiv_{l(k^{\leq n}-1)} \mathbf{B}, \overline{b} \implies \mathbf{A}, \overline{a} \sim_n \mathbf{B}, \overline{b}, \tag{3}$$

$$\mathbf{A}, \overline{a} \sim_{t_k(n)} \mathbf{B}, \overline{b} \implies \mathbf{A}, \overline{a} \equiv_n \mathbf{B}, \overline{b}, \tag{4}$$

where $t_k(0) = 0$, $t_k(n+1) = k^{\leq t_k(n)+1} + t_k(n) + 1$.

Proof: Corollary 2.7 gives (3), hence it suffices to establish (4). Clearly

$$\mathbf{A}, \overline{a} \sim_{t_k(n)} \mathbf{B}, \overline{b} \implies \mathbf{A}, \overline{a} \equiv_0 \mathbf{B}, \overline{b},$$

and by Lemma 2.9 and the definition of $t_k(n+1)$,

$$\begin{aligned}\mathbf{A}, \overline{a} \sim_{t_k(n+1)} \mathbf{B}, \overline{b} \implies &\forall c \in A\, \exists d \in B\, (\mathbf{A}, \overline{a}, c \sim_{t_k(n)} \mathbf{B}, \overline{b}, d)\\ &\wedge \forall d \in B\, \exists c \in A\, (\mathbf{A}, \overline{a}, c \sim_{t_k(n)} \mathbf{B}, \overline{b}, d).\end{aligned}$$

Thus, if $\mathbf{A}, \overline{a} \sim_{t_k(n)} \mathbf{B}, \overline{b}$, then Duplicator has a winning strategy in the n-round Ehrenfeucht–Fraïssé game for $\langle \mathbf{A}, \overline{a} \rangle$ and $\langle \mathbf{B}, \overline{b} \rangle$, which implies $\mathbf{A}, \overline{a} \equiv_n \mathbf{B}, \overline{b}$. □

Theorem 2.11 *The theory* S_k *is complete for each* $k \in \omega$. *Consequently,* $S_k = \mathrm{Th}(\mathbf{H}_k)$, *and* S_k *is decidable.*

Proof: Applying Theorem 2.10 with $l = 0$, we see that any two models of S_k are elementarily equivalent, thus S_k is complete. Being a complete recursively axiomatized theory, it is decidable. □

In order to clarify the numerical content of Theorem 2.10, let us give bounds on t_k using better known functions.

Definition 2.12 The iterated exponential function 2^x_n is defined by $2^x_0 = x$ and $2^x_{n+1} = 2^{2^x_n}$. Unless stated otherwise, $\log n$ denotes logarithm to base 2.

Proposition 2.13 *We have* $t_1(n) = 3(2^n - 1)$. *For* $k \geq 2$ *and* $n \geq 1$,

$$t_k(n) \leq 2^{c_k}_{n-1}, \tag{5}$$

where $c_k = (k+3)\log k + \log\log k + 2$.

Proof: The expression $t_1(n) = 3(2^n - 1)$ follows by induction on n from the defining recurrence, which simplifies to $t_1(0) = 0$, $t_1(n+1) = 2t_1(n) + 3$. For $k \geq 2$, we put $f(x) = k^{\leq x+1} + x + 1$ and $h(x) = (x+1)\log k + \log\log k + 2$. We want to show

$$f^{(n)}(x) \leq 2^{h(x)}_n \tag{6}$$

for all $x \geq 0$ and $n \in \omega$, which gives (5) using $t_k(n) = f^{(n)}(0) = f^{(n-1)}(k+2)$.

Now, using the monotonicity of 2^x_n in x, (6) follows by induction on n from the inequalities $x \leq h(x)$ (which is obvious) and $h(f(x)) \leq 2^{h(x)}$, hence it suffices to prove the latter; unwinding the definitions, we need to show that

$$(k^{\leq x+1} + x + 2) \log k + \log\log k + 2 \leq 4k^{x+1} \log k = 2^{(x+1)\log k + \log\log k + 2}. \tag{7}$$

This follows from the inequalities

$$k^{\leq x+1} \leq \frac{k}{k-1} k^{x+1} \leq 2k^{x+1},$$
$$x + 2 \leq 2^{x+1} \leq k^{x+1},$$
$$\log\log k + 2 \leq 2^{\log\log k + 1} \leq k \log k,$$

which are easy to verify, using the fact that $2^x \geq x + 1$ for all $x \geq 1$. □

Remark 2.14 For $k = 1$, the bound $t_1(n) = 3(2^n - 1)$ from Theorem 2.10 can be improved to $2^n - 1$, because in this case Lemma 2.9 holds with the conclusion strengthened to $\mathbf{A}, \overline{a}, c \sim_n \mathbf{B}, \overline{b}, d$. Moreover, one can also prove a matching improvement to Corollary 2.7 to obtain the exact characterization

$$\mathbf{A}, \overline{a} \equiv_n \mathbf{B}, \overline{b} \iff \mathbf{A}, \overline{a} \sim_{2^n - 1} \mathbf{B}, \overline{b},$$

using the fact that there are definitions of quantifier rank n of $y = \{x\}^t$ for each $t \leq 2^n$ and of $y = \{\varnothing\}^t$ for each $t \leq 2^n - 2$. We leave the details to an interested reader.

Apart from the completeness and decidability of S_k, Theorem 2.10 implies a quantifier elimination result for S_k:

Theorem 2.15 *Let $k \in \omega$. Then every formula is equivalent to a Boolean combination of bounded existential formulas over S_k.*

Proof: Let $\varphi(\overline{x})$ be a formula. By Theorem 2.10, there exists n such that

$$\mathbf{A}, \overline{a} \sim_n \mathbf{B}, \overline{b} \implies \big(\mathbf{A} \vDash \varphi(\overline{a}) \iff \mathbf{B} \vDash \varphi(\overline{b})\big)$$

for any $\mathbf{A}, \mathbf{B} \vDash S_k$ and $\overline{a} \in A$, $\overline{b} \in B$. There are only finitely many isomorphism types of structures of the form $\mathbf{tc}^{\mathbf{A}}_n(\overline{a})$, thus there is a finite list $\{\langle \mathbf{A}^i, \overline{a}^i \rangle : i < m\}$ such that

$$\mathbf{A} \vDash \varphi(\overline{a}) \iff \exists i < m\, (\mathbf{A}, \overline{a} \sim_n \mathbf{A}^i, \overline{a}^i).$$

Then

$$S_k \vdash \varphi(\overline{x}) \leftrightarrow \bigvee_{i<m} \varphi_{\overline{a}^i, n}(\overline{x}),$$

where $\varphi_{\overline{a}^i, n}(\overline{x})$ is as in Lemma 2.6, which makes the right-hand side a Boolean combination of bounded existential formulas. □

Remark 2.16 If we expand the language with the predicates $y = \varnothing$ and $y = \{x_0, \dots, x_{k-1}\}$ (which have bounded universal definitions in the original language), every formula is equivalent both to a bounded existential formula and to a bounded universal formula. To see this, note that an embedding $f\colon \mathbf{tc}_n^{\mathbf{A}}(\overline{a}) \to \mathbf{tc}_n^{\mathbf{B}}(\overline{b})$ in the expanded language has to be an isomorphism, as $f[\mathrm{tc}_n^{\mathbf{A}}(\overline{a})] = \mathrm{tc}_n^{\mathbf{B}}(\overline{b})$ by Lemma 2.4. It follows that if we take θ in the proof of Lemma 2.6 to be the diagram in the expanded language, then it suffices to put $\varphi_{\overline{a},n} = \psi$.

If $k = 1$, the $y = \{x\}$ predicate is redundant, as it is equivalent to $x \in y$. Moreover, S_1 has full quantifier elimination in a language with function symbols $\varnothing$ and $\{x\}$, as $\langle H_1, \varnothing, \{x\}\rangle \simeq \langle \mathbb{N}, 0, S\rangle$.

As we learned from Albert Visser, it is an interesting problem whether there exists a *finitely axiomatized* consistent decidable theory with a pairing function. We observe that our theories do not cut the mustard, though we postpone the (albeit simple) proof to the next section, where the relevant construction will be used in a more substantial way:

Proposition 2.17 *S_k is not finitely axiomatizable for any $k > 0$.*

Proof: See Corollary 3.5. $\square$

Remark 2.18 The axioms (C_n) of S_k express the acyclicity of $\in$. More generally, since $\in$ is well founded, $\mathbf{H}_k$ satisfies the $\in$-*induction* schema

$$\forall x\, \big(\forall y \in x\, \varphi(y) \to \varphi(x)\big) \to \forall x\, \varphi(x)$$

(where φ is any formula, possibly with parameters), of which each (C_n) axiom is a special case. By Theorem 2.11, the full $\in$-induction schema is equivalent to its instances $\{(\mathrm{C}_n) : n \geq 1\}$ over the remaining axioms of S_k; there does not seem to be an easy direct proof of this fact.

The axiom of foundation (regularity) as commonly formulated in ZF,

$$x \neq \varnothing \to \exists y \in x\, \forall z\, \neg(z \in x \wedge z \in y),$$

is strictly weaker: using the fact that $|x| \leq k$ by (B_k), it is easily seen to be equivalent to $\{(\mathrm{C}_n) : 1 \leq n \leq k\}$.

We end this section with a basic model-theoretic classification of the S_k theories.

Definition 2.19 Let κ be an infinite cardinal. A theory T is κ-*stable* if for every $\mathbf{M} \vDash T$ and $A \subseteq M$ of size $|A| \leq \kappa$, there are at most κ (complete) 1-types of $\mathbf{M}$ over A. We say that T is *stable* if it is κ-stable for some $\kappa \geq \|T\|$, and it is *superstable* if there is κ_0 such that T is κ-stable for all $\kappa \geq \kappa_0$.

As is well known, the theory S_1—definitionally equivalent to $\mathrm{Th}(\mathbb{N}, 0, S)$—is uncountably categorical, and therefore κ-stable for all $\kappa \geq \omega$. In contrast to that, it is easy to see that no consistent theory with pairing (even non-functional) can be superstable, as there are always at least $|A|^\omega$ different types over A; thus, the result below is the best possible for $k \geq 2$.

Proposition 2.20 *For each $k \geq 2$, the theory S_k is stable.*

Proof: Let $\mathbf{M} \vDash S_k$, and $A \subseteq M$ be such that $|A| \leq \kappa$. By replacing A with $\mathrm{tc}^{\mathbf{M}}(A)$ (which has the same cardinality) if necessary, we may assume $\mathrm{tc}^{\mathbf{M}}(A) = A$. By Theorem 2.10, 1-types over A correspond to isomorphism types of $\mathbf{tc}^{\mathbf{N}}(A, c)$ for $\mathbf{N} \succeq \mathbf{M}$, $c \in N$; since the structure on A is fixed, these are determined by isomorphism types of $\mathbf{tc}^{\mathbf{N}}(c)$ expanded with constants a for all $a \in A \cap \mathrm{tc}^{\mathbf{N}}(c)$. In other words, these structures are certain countable pointed directed graphs endowed with a partial vertex labelling with labels from A. Thus, the number of types is at most $2^\omega |A|^\omega \leq \kappa^\omega$, and consequently, S_k is κ-stable whenever $\kappa = \kappa^\omega$. $\square$

We remark that theories of locally free algebras (including acyclic pairing functions) are also stable; further model-theoretic properties of acyclic pairing functions were investigated by Bouscaren and Poizat [2].

3 Computational complexity

The proof of Theorem 2.11 does not give any bound on the computational complexity of S_k, but as we will see in this section, we can actually find reasonably tight upper and lower bounds on the complexity. Recall that there is a general lower bound due to Ferrante and Rackoff [7]:

Theorem 3.1 *Let T be a consistent theory with a pairing function. Then every language $L \in \mathrm{DTIME}(2^0_{O(n)})$ has a linearly-bounded polynomial-time reduction to T. Consequently, there exists $\gamma > 0$ such that every decision procedure for T takes time at least $2^0_{\gamma n}$ for infinitely many input lengths n.* $\square$

A few remarks are in order. First, the result is stated in [7] for theories of a pairing *function*, but it is straightforward to adapt the argument to theories with a non-functional pairing predicate. The constant γ only depends on the defining formula for the pairing predicate, otherwise it is independent of T. Second, the result is quite robust across models of computation and complexity measures: it applies equally well to time or space, on deterministic, nondeterministic, or alternating Turing machines, etc. The reason is that all these measures are equivalent up to an exponential or two, and this difference is drowned by the overall complexity: say, $\mathrm{ASPACE}(2^0_{\gamma n}) \subseteq \mathrm{DTIME}(2^0_{\gamma n + O(1)})$.

In Theorem 3.1, the length of input is officially measured as the number of letters when the formula is written as a word over a finite alphabet (thus a variable x_i takes length $O(\log i)$), but a fortiori the bound also holds when we measure the input by the number of symbols (quantifiers, connectives, variables, predicate and function symbols); as we will see, the bound is tight in both regimes (the explanation is that the formulas used in the lower bound reuse just $O(1)$ distinct variables over and over). We will state upper bounds in terms of the number of symbols, which is more intuitive, and makes the upper bounds stronger.

Corollary 3.2 *There exists $\gamma > 0$ such that every decision procedure for any consistent extension of VS_2 has complexity at least $2^0_{\gamma n}$ for infinitely many input lengths n. In particular, this applies to the theories S_k for $k \geq 2$.* $\square$

We aim to show that the bound on the complexity of S_k from Corollary 3.2 is optimal up to the value of γ. The basic idea is that using Theorem 2.10, we can represent tuples from an unspecified model of S_k by finite objects of bounded size (namely, isomorphism types of $\mathbf{tc}_m(\overline{a})$ for sufficiently large m) that carry enough information to determine the truth of $\varphi(\overline{a})$ for a given formula φ. First, we need an internal description of structures of the form $\mathbf{tc}^{\mathbf{A}}_m(\overline{a})$ so that we can efficiently recognize them.

Definition 3.3 Consider a (possibly empty) structure $\mathbf{T} = \langle T, \in^{\mathbf{T}}, \overline{a}\rangle$, where $\mathrm{lh}(\overline{a}) = l$. We regard $\mathbf{T}$ as a directed graph such that there is an edge $x \to y$ iff $y \in^{\mathbf{T}} x$. We say that $\mathbf{T}$ is a $\mathrm{tc}^k_m(l)$*-structure* if it satisfies the following conditions:

- $\mathbf{T}$ is a directed acyclic graph with all nodes of out-degree $\le k$.
- Every node $t \in T$ is reachable from some a_i, $i < l$, in at most m steps.
- Let $U \subseteq T$ denote the set of nodes $u \in T$ such that u is reachable from some a_i in $< m$ steps, or u has out-degree k. Then $\mathbf{T}$ is extensional w.r.t. U: i.e., for every distinct $u, u' \in U$, there is $t \in T$ such that $t \in^{\mathbf{T}} u$ and $t \notin^{\mathbf{T}} u'$, or vice versa.

Lemma 3.4 *For any structure* $\mathbf{T} = \langle T, \in^{\mathbf{T}}, \overline{a}\rangle$ *with* $\mathrm{lh}(\overline{a}) = l$*, the following are equivalent:*

(i) $\mathbf{T}$ *is a* $\mathrm{tc}^k_m(l)$*-structure.*

(ii) $\mathbf{T}$ *embeds in a model* $\mathbf{A} \vDash S_k$ *in such a way that* $\mathbf{T} = \mathbf{tc}^{\mathbf{A}}_m(\overline{a})$*.*

Proof: (ii) $\to$ (i) is clear, using the observation that $u^{\mathbf{A}} \subseteq \mathrm{tc}^{\mathbf{A}}_m(\overline{a})$ for every u from the set

$$U = \mathrm{tc}^{\mathbf{A}}_{m-1}(\overline{a}) \cup \bigl\{u \in \mathrm{tc}^{\mathbf{A}}_m(\overline{a}) : |u^{\mathbf{A}} \cap \mathrm{tc}^{\mathbf{A}}_m(\overline{a})| = k\bigr\}.$$

(i) $\to$ (ii): Let U be as in Definition 3.3. We first extend $\mathbf{T}$ to a model $\mathbf{A}_0 = \langle A_0, \in^{\mathbf{A}_0}\rangle$ by adding an infinite descending chain below each $u \in T \smallsetminus U$; formally, $A_0 = T \mathbin{\dot\cup} \bigl((T \smallsetminus U) \times \omega\bigr)$, with

$$\in^{\mathbf{A}_0} = \in^{\mathbf{T}} \cup \bigl\{\bigl\langle\langle u, 0\rangle, u\bigr\rangle, \bigl\langle\langle u, n+1\rangle, \langle u, n\rangle\bigr\rangle : u \in T \smallsetminus U, n \in \omega\bigr\}.$$

Notice that no element of $A_0 \smallsetminus T$ is reachable in $\le m$ steps from $\overline{a}$. Since each $u \in T \smallsetminus U$ has strictly less than k elements in $\mathbf{T}$, all $u \in A_0$ have at most k elements in $\mathbf{A}_0$. Moreover, the structure is still acyclic, and the added chains ensure that it is extensional; i.e., $\mathbf{A}_0$ satisfies the axioms (E), (B_k), and (C_n) for all $n \ge 1$.

In order to satisfy axioms (V_0) and (V_k) as well, we inductively add to $\mathbf{A}_0$ all the missing subsets of size at most k: i.e., we define $\mathbf{A}_i = \langle A_i, \in^{\mathbf{A}_i}\rangle$ by induction on $i \in \omega$ as

$$\begin{aligned} A_{i+1} &= A_i \mathbin{\dot\cup} \{x \subseteq A_i : |x| \le k, \forall u \in A_i\ u^{\mathbf{A}_i} \ne x\}, \\ \in^{\mathbf{A}_{i+1}} &= \in^{\mathbf{A}_i} \cup \{\langle u, x\rangle : x \in A_{i+1} \smallsetminus A_i, u \in x\}, \end{aligned}$$

and we let $\mathbf{A} = \langle A, \in^{\mathbf{A}}\rangle$ be the union of the chain:

$$A = \bigcup_{i \in \omega} A_i, \qquad \in^{\mathbf{A}} = \bigcup_{i \in \omega} \in^{\mathbf{A}_i}.$$

By construction, $\mathbf{A} \vDash S_k$ and $\mathrm{tc}^{\mathbf{A}}_m(\overline{a}) = \mathbf{T}$. $\square$

Incidentally, the construction from Lemma 3.4 can be used to prove Proposition 2.17:

Corollary 3.5 *S_k is not finitely axiomatizable for any $k > 0$.*

Proof: Since any finite set of consequences of S_k is provable from a finite subset of the axiomatization of S_k in Definition 2.1, it suffices to show that for every $n \geq 1$, there is a model $\mathbf{A}$ satisfying (V_0), (V_k), (E), (B_k), (C_i) for $1 \leq i < n$, and $\neg$(C_n). Let $\mathbf{A}_0$ be an n-cycle, and build $\mathbf{A}$ from $\mathbf{A}_0$ as in the proof of Lemma 3.4. $\square$

Definition 3.6 Let $k, l, m \in \omega$. If $\mathbf{T} = \langle T, \in^{\mathbf{T}}, \overline{a} \rangle$ is a $\mathrm{tc}^k_m(l)$-structure, and $\varphi(\overline{x})$ a formula such that $l = \mathrm{lh}(\overline{x})$ and $m \geq t_k(\mathrm{rk}(\varphi))$, we write $\mathbf{T} \vDash_{S_k} \varphi(\overline{a})$ if $\mathbf{A} \vDash \varphi(\overline{a})$, where $\mathbf{A} \vDash S_k$ is such that $\mathbf{T} = \mathrm{tc}^{\mathbf{A}}_m(\overline{a})$. (Such an $\mathbf{A}$ exists by Lemma 3.4, and the definition is independent of the choice of $\mathbf{A}$ by Theorem 2.10.)

If $\mathbf{T} = \langle T, \in^{\mathbf{T}}, \overline{a} \rangle$ is a $\mathrm{tc}^k_m(l)$-structure, $m' \leq m$, $l' \leq l$, and $\overline{a}'$ is a subsequence of $\overline{a}$ of length l', let $\mathbf{tc}^{\mathbf{T}}_{m'}(\overline{a}')$ denote the $\mathrm{tc}^k_{m'}(l')$-structure $\langle T', \in^{\mathbf{T}} \cap (T' \times T'), \overline{a}' \rangle$, where T' is the set of nodes of T reachable in $\leq m'$ steps from $\overline{a}'$. (This coincides with $\mathbf{tc}^{\mathbf{A}}_{m'}(\overline{a}')$ for any $\mathbf{A} \vDash S_k$ such that $\mathbf{T} = \mathbf{tc}^{\mathbf{A}}_m(\overline{a})$.)

For testing the truth of quantified formulas in $\mathbf{T}$, we will need to be able to efficiently recognize when a $\mathrm{tc}^k_m(l)$-structure and a $\mathrm{tc}^k_{m'}(l')$-structure are compatible in that they can be jointly embedded in a model of S_k. This is accomplished in the next lemma; note that $\overline{a}$ and $\overline{a}'$ need not be disjoint (in fact, the intended use case is that $\overline{a}'$ extends $\overline{a}$).

Lemma 3.7 *Let $\mathbf{T} = \langle T, \in^{\mathbf{T}}, \overline{a}, \overline{a}' \rangle$, $l = \mathrm{lh}(\overline{a})$, $l' = \mathrm{lh}(\overline{a}')$, and $k, m, m' \geq 0$. Then the following are equivalent.*

(i) $\mathbf{T}$ *embeds in a model $\mathbf{A} \vDash S_k$ in such a way that $T = \mathrm{tc}^{\mathbf{A}}_m(\overline{a}) \cup \mathrm{tc}^{\mathbf{A}}_{m'}(\overline{a}')$.*

(ii) *The following conditions hold:*

- $\mathbf{T}$ *is a directed acyclic graph with all nodes of out-degree $\leq k$.*
- *Every node $t \in T$ is reachable from some a_i, $i < l$, in at most m steps, or from some a'_i, $i < l'$, in at most m' steps.*
- $\mathbf{T}$ *is extensional w.r.t. U, where U denotes the set of nodes $u \in T$ such that u is reachable from some a_i in $< m$ steps, or from some a'_i in $< m'$ steps, or u has out-degree k.*

Proof: Just like the proof of Lemma 3.4. $\square$

If $\mathbf{T}$ satisfies the conditions of Lemma 3.7, the $\mathrm{tc}^k_m(l)$-structure $\mathbf{tc}^{\mathbf{T}}_m(\overline{a})$ and the $\mathrm{tc}^k_{m'}(l')$-structure $\mathbf{tc}^{\mathbf{T}}_{m'}(\overline{a}')$ are called *compatible*. Note that $\mathbf{T}$ is uniquely determined by $\mathbf{tc}^{\mathbf{T}}_m(\overline{a})$ and $\mathbf{tc}^{\mathbf{T}}_{m'}(\overline{a}')$, being their union. We stress that compatibility is not defined "up to isomorphism"; the two structures have to be presented in such a way that elements of their intersection inside $\mathbf{T}$ are represented literally the same in both.

We consider the recursive algorithm S_k-Sat in Fig. 1. (We are primarily interested in the case where k is a constant, but the algorithm actually works uniformly even if k is given as part of the input.)

function S_k-Sat$(\mathbf{T}, \varphi) \in \{0, 1\}$
input: $\mathrm{tc}^k_m(l)$-structure $\mathbf{T} = \langle T, \in^{\mathbf{T}}, \overline{a} \rangle$, formula $\varphi(\overline{x})$,
    where $l = \mathrm{lh}(\overline{x})$, $m \geq t_k(\mathrm{rk}(\varphi))$
1 **if** φ is atomic **then return** $\mathbf{T} \vDash \varphi(\overline{a})$
2 **if** $\varphi = \neg\varphi_0$ **then return** $\neg S_k$-Sat$(\mathbf{T}, \varphi_0)$
3 **if** $\varphi = \varphi_0 \vee \varphi_1$ **then return** S_k-Sat$(\mathbf{T}, \varphi_0) \vee S_k$-Sat$(\mathbf{T}, \varphi_1)$
4 **if** $\varphi = \varphi_0 \wedge \varphi_1$ **then return** S_k-Sat$(\mathbf{T}, \varphi_0) \wedge S_k$-Sat$(\mathbf{T}, \varphi_1)$
5 **if** $\varphi = \exists y\, \varphi_0(\overline{x}, y)$ **then**:
6   **for each** $\mathrm{tc}^k_{t_k(\mathrm{rk}(\varphi_0))}(l+1)$-structure $\mathbf{T}' = \langle T', \in^{\mathbf{T}'}, \overline{a}, c \rangle$ **do**:
7   **if** $\mathbf{T}'$ is compatible with $\mathbf{T}$ and S_k-Sat$(\mathbf{T}', \varphi_0) = 1$ **then return** 1
8   **return** 0
9 **if** $\varphi = \forall y\, \varphi_0(\overline{x}, y)$ **then**:
10   **for each** $\mathrm{tc}^k_{t_k(\mathrm{rk}(\varphi_0))}(l+1)$-structure $\mathbf{T}' = \langle T', \in^{\mathbf{T}'}, \overline{a}, c \rangle$ **do**:
11   **if** $\mathbf{T}'$ is compatible with $\mathbf{T}$ and S_k-Sat$(\mathbf{T}', \varphi_0) = 0$ **then return** 0
12   **return** 1

Figure 1: An algorithm for $\mathbf{T} \vDash_{S_k} \varphi(\overline{a})$.

Lemma 3.8 *Given a* $\mathrm{tc}^k_m(l)$*-structure* $\mathbf{T} = \langle T, \in^{\mathbf{T}}, \overline{a} \rangle$ *and a formula* $\varphi(\overline{x})$ *such that* $l = \mathrm{lh}(\overline{x})$ *and* $m \geq t_k(\mathrm{rk}(\varphi))$, S_k-Sat$(\mathbf{T}, \varphi) = 1$ *if and only if* $\mathbf{T} \vDash_{S_k} \varphi(\overline{a})$.

Proof: By induction on the complexity of φ. The only nontrivial cases are for the quantifiers. We will give the proof for $\varphi(\overline{x}) = \exists y\, \varphi_0(\overline{x}, y)$; the argument for $\forall y\, \varphi_0(\overline{x}, y)$ is dual.

On the one hand, assume that $\mathbf{T} \vDash_{S_k} \varphi(\overline{a})$; i.e., $\mathbf{A} \vDash \varphi(\overline{a})$, where we fix $\mathbf{A} \vDash S_k$ such that $\mathbf{T} = \mathbf{tc}^{\mathbf{A}}_m(\overline{a})$. Let $c \in A$ be such that $\mathbf{A} \vDash \varphi_0(\overline{a}, c)$, and put $m' = t_k(\mathrm{rk}(\varphi_0))$. Then $\mathbf{T}' = \mathbf{tc}^{\mathbf{A}}_{m'}(\overline{a}, c)$ is a $\mathrm{tc}^k_{m'}(l+1)$-structure compatible with $\mathbf{T}$, and $\mathbf{T}' \vDash_{S_k} \varphi_0(\overline{a}, c)$, hence S_k-Sat$(\mathbf{T}', \varphi_0) = 1$ by the induction hypothesis. Thus, S_k-Sat$(\mathbf{T}, \varphi)$ returns 1 on line 7.

On the other hand, assume that S_k-Sat$(\mathbf{T}, \varphi) = 1$, thus there is a $\mathrm{tc}^k_{m'}(l+1)$-structure $\mathbf{T}' = \langle T', \in^{\mathbf{T}'}, \overline{a}, c \rangle$ compatible with $\mathbf{T}$ such that S_k-Sat$(\mathbf{T}', \varphi_0) = 1$. By compatibility, there is a model $\mathbf{A} \vDash S_k$ such that $\mathbf{T} = \mathbf{tc}^{\mathbf{A}}_m(\overline{a})$ and $\mathbf{T}' = \mathbf{tc}^{\mathbf{A}}_{m'}(\overline{a}, c)$. By the induction hypothesis, $\mathbf{T}' \vDash_{S_k} \varphi_0(\overline{a}, c)$, which means $\mathbf{A} \vDash \varphi_0(\overline{a}, c)$ and $\mathbf{A} \vDash \varphi(\overline{a})$. Thus, $\mathbf{T} \vDash_{S_k} \varphi(\overline{a})$. $\square$

Theorem 3.9 *Let* $k \geq 2$. *Given a sentence* φ *with* n *symbols, we can decide whether* $S_k \vdash \varphi$ *in time* $2^{c_k}_{(n+1)/4}$ *for sufficiently large* n*, where* c_k *is the constant from Proposition 2.13.*

Proof: We have $S_k \vdash \varphi$ iff $\varnothing \vDash_{S_k} \varphi$ iff S_k-Sat$(\varnothing, \varphi) = 1$ by Lemma 3.8, where $\varnothing$ is considered as a $\mathrm{tc}^k_m(0)$-structure with $m = t_k(\mathrm{rk}(\varphi))$. Rather than measuring time directly, it is easier to estimate the space requirements of S_k-Sat$(\varnothing, \varphi)$. We claim that space $O(m \log m + n)$ is sufficient.

It is easy to see that we can test whether a given $\mathbf{T}$ is a $\mathrm{tc}^k_m(l)$-structure in space linear in the size of $\mathbf{T}$; likewise for testing compatibility, or the truth of atomic formulas. Thus, the dominant cost is that for each recursive call, we need to store $O(1)$ bits describing where the call was made, and for the quantifier cases, the structure $\mathbf{T}'$. The former add up to space

$O(n)$, as the recursion depth is at most n. The latter are dominated by the size of $\mathbf{T}'$ in the top-most quantifier calls, where it has $s \leq k^{\leq t_k(\mathrm{rk}(\varphi_0))} \leq m/k$ elements (in subsequent calls, the structures become exponentially smaller, hence their space requirements are negligible in comparison). Since $\mathbf{T}'$ is a directed graph with out-degree at most k, it can be described by a list of edges using $O(ks \log s) = O(m \log m)$ bits; this gives total space $O(m \log m + n)$. As long as m dominates n (which will be the case for our bounds on m below), this means the algorithm works in space $O(m \log m)$, and therefore in time $m^{O(m)}$.

In order to bound m in terms of n, we first bound $r = \mathrm{rk}(\varphi)$. Obviously, $r \leq n$, but we may do a bit better as follows. By preprocessing φ if necessary, we may assume that there are no dummy quantifiers in φ. Then each quantified variable occurs also in an atomic formula; since only two variables occur in a single atomic formula, it follows that the formula has $\geq r/2$ atomic subformulas (of 3 symbols each), and consequently $\geq r/2 - 1$ binary connectives. Since every quantifier takes two symbols by itself, we see that $n \geq 4r - 1$, i.e., $r \leq (n+1)/4$.

By Proposition 2.13, $m \leq 2^{c_k}_{r-1}$, where $r = \mathrm{rk}(\varphi)$. Thus, $m \leq 2^{c_k}_{(n-3)/4}$. Since this grows much faster than n, we obtain that the algorithm works in space $O(2^{c_k}_{(n-3)/4} \log 2^{c_k}_{(n-3)/4})$. In fact, it is easy to check that there is enough leeway in the bound from Proposition 2.13 so that for any constant C, $Ct_k(r) \log t_k(r) \leq 2^{c_k}_{r-1}$ for large enough r. Thus, for large enough n, the algorithm works in space $2^{c_k}_{(n-3)/4}$, and in time $2^{c_k}_{(n+1)/4}$. □

The main virtue of Theorem 3.9 is that it provides an upper bound on the complexity of S_k that matches the lower bound from Theorem 3.1 up to the value of γ, and to that end it is stated so that the bound only depends on n (and k, which is considered to be constant), not other parameters. On the flip side, this simplicity means that it vastly overestimates the needed complexity for many classes of formulas.

It is clear from the proof that the height of the tower of exponentials in the bound is actually controlled by the quantifier rank rather than the length of the sentence. Even better, we will show below that it only depends on the number of quantifier *alternations*.

For simplicity, we will formulate the result for sentences in prenex normal form. Recall that a formula is $\exists_n$ if it is in prenex normal form, and the quantifier prefix consists of n alternating (possibly empty) blocks of quantifiers, where the first block is existential. The definition of $\forall_n$ formulas is dual. Let us first generalize Lemma 2.9 and Theorem 2.10 to handle blocks of quantifiers.

Lemma 3.10 *Let $\mathbf{A}, \mathbf{B} \vDash S_k$, $\overline{a} \in A$, $\overline{b} \in B$, $l = \mathrm{lh}(\overline{a}) = \mathrm{lh}(\overline{b})$, and $n, q > 0$. If $\mathbf{A}, \overline{a} \sim_{q\,k^{\leq n}+n} \mathbf{B}, \overline{b}$, then for every q-tuple $\overline{c} \in A$, there exists a q-tuple $\overline{d} \in B$ such that $\mathbf{A}, \overline{a}, \overline{c} \sim_{n-1} \mathbf{B}, \overline{b}, \overline{d}$.*

Proof: The proof of Lemma 2.9 works literally the same with $\overline{c}$ in place of c, and $q\,k^{\leq n}$ in place of $k^{\leq n}$. In particular, the quantity $k^{\leq n}$ only enters the proof through the bound $|\mathrm{tc}^{\mathbf{A}}_n(c)| \leq k^{\leq n}$, which is now replaced with $|\mathrm{tc}^{\mathbf{A}}_n(\overline{c})| \leq q\,k^{\leq n}$. □

Theorem 3.11 *Let $\mathbf{A}, \mathbf{B} \vDash S_k$, $\overline{a} \in A$, $\overline{b} \in B$, and $\mathrm{lh}(\overline{a}) = \mathrm{lh}(\overline{b})$. For any $n, q \in \omega$, define $t_k(n, q)$ by $t_k(0, q) = 0$, $t_k(n+1, q) = q\,k^{\leq t_k(n,q)+1} + t_k(n, q) + 1$. Let $\varphi(\overline{x})$ be an $\exists_n$ formula with each quantifier block of length at most q. Then*

$$\mathbf{A}, \overline{a} \sim_{t_k(n,q)} \mathbf{B}, \overline{b} \implies \big(\mathbf{A} \vDash \varphi(\overline{a}) \iff \mathbf{B} \vDash \varphi(\overline{b})\big).$$

function S_k-BSat$(\mathbf{T}, \varphi) \in \{0,1\}$
input: $\mathrm{tc}^k_m(l)$-structure $\mathbf{T} = \langle T, \in^{\mathbf{T}}, \overline{a}\rangle$, $\exists_r$ or $\forall_r$ formula $\varphi(\overline{x})$,
where $l = \mathrm{lh}(\overline{x})$, $m \geq t_k(r,q)$, $q =$ maximal quantifier block size in φ
if $r = 0$ **then return** $\mathbf{T} \vDash \varphi(\overline{a})$
if $\varphi = \exists \overline{y}\, \varphi_0(\overline{x}, \overline{y})$, $\varphi_0 \in \forall_{r-1}$, $l_0 = \mathrm{lh}(\overline{y})$ **then**:
    for each $\mathrm{tc}^k_{t_k(r-1,q)}(l+l_0)$-structure $\mathbf{T}' = \langle T', \in^{\mathbf{T}'}, \overline{a}, \overline{c}\rangle$ **do**:
        if $\mathbf{T}'$ is compatible with $\mathbf{T}$ and S_k-BSat$(\mathbf{T}', \varphi_0) = 1$ **then return** 1
    return 0
if $\varphi = \forall \overline{y}\, \varphi_0(\overline{x}, \overline{y})$, $\varphi_0 \in \exists_{r-1}$, $l_0 = \mathrm{lh}(\overline{y})$ **then**:
    for each $\mathrm{tc}^k_{t_k(r-1,q)}(l+l_0)$-structure $\mathbf{T}' = \langle T', \in^{\mathbf{T}'}, \overline{a}, \overline{c}\rangle$ **do**:
        if $\mathbf{T}'$ is compatible with $\mathbf{T}$ and S_k-BSat$(\mathbf{T}', \varphi_0) = 0$ **then return** 0
    return 1

Figure 2: A block-wise algorithm for $\mathbf{T} \vDash_{S_k} \varphi(\overline{a})$.

Proof: By induction on n, using Lemma 3.10. □

Lemma 3.12 *For any $k \geq 2$ and $n, q \geq 1$, we have*

$$t_k(n,q) \leq 2_{n-1}^{(q(k+1)+2)\log k + \log\log k + \log q + 2} \leq 2_{n-1}^{4qk\log k}.$$

Proof: Similar to the proof of Proposition 2.13, with $(x+1)\log k + \log\log k + \log q + 2$ in place of $h(x)$, using the inequality

$$(q\, k^{\leq x+1} + x + 2)\log k + \log\log k + \log q + 2 \leq 4q\, k^{x+1}\log k,$$

which can be proved in the same way as (7). □

Theorem 3.13 *Given a sentence φ in prenex normal form and $k \geq 2$, we can decide whether $S_k \vdash \varphi$ in $\mathrm{NTIME}\big(t_k(r,q)^{O(t_k(r,q))} n^{O(1)}\big)$, where n is the length of φ, r is such that φ is $\exists_{r+1}$, and q is the maximal length of a quantifier block in φ. This is $\mathrm{NTIME}(n^{O(1)})$ for $r = 0$, $\mathrm{NTIME}\big((kq)^{O(kq)} n^{O(1)}\big)$ for $r = 1$, and $\mathrm{NTIME}\big(2_r^{O(qk\log k)} n^{O(1)}\big)$ for $r \geq 2$.*

Proof: Write $\varphi = \exists \overline{x}\, \psi(\overline{x})$, where ψ is $\forall_r$. Put $m = t_k(r,q)$ and $l = \mathrm{lh}(\overline{x}) \leq q$. In order to test $S_k \vdash \varphi$, we nondeterministically guess a $\mathrm{tc}^k_m(l)$-structure $\mathbf{T} = \langle T, \in^{\mathbf{T}}, \overline{a}\rangle$, and verify $\mathbf{T} \vDash_{S_k} \psi(\overline{a})$ using the algorithm S_k-BSat$(\mathbf{T}, \psi)$ from Fig. 2. Note that $s = |T| \leq l\, k^{\leq m} \leq \frac{1}{k} t_k(r+1, q)$, hence the bit-size of $\mathbf{T}$ is $O(ks\log s) = O\big(t_k(r+1,q)\log t_k(r+1,q)\big)$, and we can check that $\mathbf{T}$ is a $\mathrm{tc}^k_m(l)$-structure in time polynomial in $t_k(r+1,q) = q\, k^{O(t_k(r,q))}$. (For $r = 0$, we have $s \leq l \leq n$, thus $\mathbf{T}$ can be represented with $O(n^2)$ bits using an adjacency matrix, and then we can check that $\mathbf{T}$ is a $\mathrm{tc}^k_0(l)$-structure in time $n^{O(1)}$ independent of k: if $k > n$, we only need to check that $T = \{\overline{a}\}$ and $\in^{\mathbf{T}}$ is acyclic.)

We claim that S_k-BSat$(\mathbf{T}, \psi)$, and thus the whole test, works in time polynomial in n and $t_k(r,q)^{t_k(r,q)}$. In the top-level iteration, the structures $\mathbf{T}'$ have sizes up to $(l + l_0)k^{\leq t_k(r-1,q)} \leq 2q\, k^{\leq t_k(r-1,q)} \leq \frac{2}{k} t_k(r,q)$, and can be described using $O\big(t_k(r,q)\log t_k(r,q)\big)$ bits. Thus, the loop on lines 7–8 goes through $\exp\bigl(O\bigl(t_k(r,q)\log t_k(r,q)\bigr)\bigr)$ structures $\mathbf{T}'$; for each of them, it checks in time $t_k(r+1,q)^{O(1)}$ whether it is compatible with $\mathbf{T}$, and if so, makes a recursive call. In turn, each of these recursive calls will involve a loop over $\exp\bigl(O\bigl(t_k(r-1,q)\log t_k(r-1,q)\bigr)\bigr)$ structures, where for each of them, we do a compatibility check in time $t_k(r,q)^{O(1)}$, and a recursive call. This goes on until we get down to the quantifier-free matrix at recursion depth r; this takes time $n^{O(1)}$ to check on line 1. Thus, the total number of recursive calls is

$$\prod_{i<r} 2^{O(t_k(r-i,q)\log t_k(r-i,q))} = 2^{O\left(\sum_{i<r} t_k(r-i,q)\log t_k(r-i,q)\right)}$$
$$= 2^{O(t_k(r,q)\log t_k(r,q))} = t_k(r,q)^{O(t_k(r,q))},$$

and each takes time polynomial in $t_k(r+1,q) = q\,k^{O(t_k(r,q))}$ and n. This gives total time $t_k(r,q)^{O(t_k(r,q))}n^{O(1)}$, as claimed. (For $r=0$, this means $n^{O(1)}$; there are no recursive calls.)

For $r=1$, we have $t_k(1,q) = O(kq)$, hence the time bound is $(kq)^{O(kq)}n^{O(1)}$. For $r \geq 2$, $t_k(r,q) = 2_{r-1}^{O(qk\log k)}$ by Lemma 3.12; it is easy to show that $(2_d^x)^c \leq 2_d^{cx}$ for any $d, x \geq 1$ by induction on d, hence

$$t_k(r,q)^{O(t_k(r,q))} = 2^{t_k(r,q)^{O(1)}} = 2^{\left(2_{r-1}^{O(qk\log k)}\right)^{O(1)}} = 2^{2_{r-1}^{O(qk\log k)}} = 2_r^{O(qk\log k)},$$

which gives the time bound $2_r^{O(qk\log k)}n^{O(1)}$. □

For completeness, let us also indicate the complexity of S_k for $k=0,1$, which is essentially known from the literature.

Theorem 3.14 *S_1 is* PSPACE*-complete, and for any fixed $r \geq 1$, the $\exists_r$ fragment of S_1 is Σ_r^{P}-complete, and the $\forall_r$ fragment is Π_r^{P}-complete.*

S_0 is decidable in P*; more precisely, it is complete for* ALOGTIME $=$ *U_E-uniform* NC^1 *under* DLOGTIME *reductions.*

Proof: S_0 is the theory of a one-element structure, hence it is equivalent to propositional logic (we can decide a given sentence by removing all quantifiers, replacing atomic formulas $x_i \in x_j$ with the truth-constant 0 and $x_i = x_j$ with 1, and evaluating the resulting Boolean sentence). This is NC^1-complete by results of Buss [3].

For S_1, it is well known and easy to see that the truth of quantified Boolean sentences is reducible to any consistent first-order theory T that proves the existence of two distinct elements, making T PSPACE-hard. Moreover, the reduction takes $\exists_r$ QBF to $\exists_r$ sentences, hence T-provability of $\exists_r$ sentences is Σ_r^{P}-hard, and dually for $\forall_r$.

On the other hand, $\mathbf{H}_1$ is definitionally equivalent to $\langle \mathbb{N}, 0, S\rangle$, whose theory is known to be decidable in PSPACE: e.g., this is proved in [7] for the more general structure $\langle \mathbb{N}, <\rangle$.

Using the machinery we have already developed, this can be shown as follows. First, we have $t_1(r,q) = (2+\frac{1}{q})(q+1)^r$, hence S_1-BSat$(\varnothing,\varphi)$ decides $S_1 \vdash \varphi$ in exponential space. We can make it more space-efficient by employing a more compact representation for $\mathrm{tc}_m^1(l)$-structures $\mathbf{T} = \langle T, \in^{\mathbf{T}}, \overline{a}\rangle$. Any such structure is a disjoint union of $\in^{\mathbf{T}}$-chains, where each chain has some a_i on top, the distance between neighbouring a_i and a_j on the same chain is $\leq m+1$, and each chain ends $\leq m$ steps below the lowest a_i on the chain. We can represent this by noting for each a_i the nearest a_j below a_i on the same chain, if any, and the distance (in binary!) from a_i to a_j, or to the end of the chain. This takes $O(l \log m) = O(lr \log(q+1)) = O(n^2 \log n)$ bits if $m = t_1(r, q)$ and $l, r, q \leq n$, and we can test compatibility of $\mathrm{tc}^1_m(l)$-structures in this representation and satisfaction of quantifier-free formulas in polynomial time.

Thus, S_1-BSat modified to use this representation runs in polynomial space, placing S_1 in PSPACE. Moreover, we may view the modified S_1-BSat as an alternating polynomial-time algorithm, where the loop on lines 3–4 is replaced with a nondeterministic (existential) guess of $\mathbf{T}'$, and the loop on lines 7–8 with a co-nondeterministic (universal) guess. Then S_1-BSat$(\varnothing, \varphi)$ for an $\exists_r$ sentence φ makes $r-1$ alternations starting from an existential state, i.e., it works in Σ_r-TIME$(n^{O(1)}) = \Sigma^{\mathrm{P}}_r$, and dually for $\forall_r$ sentences. $\square$

4 Conclusion

As we have seen, the complete theory of the structure $\mathbf{H}_k$ can be described by a simple list of axioms, it is decidable, and generally tame (it has quantifier elimination down to formulas of quite a low complexity, it is stable, and its computational complexity—albeit somewhat daunting—is the lowest possible for theories with pairing). Thus, it is in many respects as nice as other known examples of decidable theories with a pairing function.

However, it has a different flavour from the previous examples, which are generally of algebraic or arithmetic nature, whereas here we have a theory of sets. In particular, the theories $\mathrm{Th}(\mathbf{H}_k)$ provide natural decidable extensions of finite fragments of the Vaught set theory *VS*, which was our original motivation.

Acknowledgements

I want to thank Albert Visser for a helpful discussion on theories of pairing, and an anonymous reviewer for useful comments. A preliminary sketch of the basic results of this paper was first reported in [8].

References

[1] Wilhelm Ackermann, *Die Widerspruchsfreiheit der allgemeinen Mengenlehre*, Mathematische Annalen 114 (1937), pp. 305–315 (in German).

[2] Elisabeth Bouscaren and Bruno Poizat, *Des belles paires aux beaux uples*, Journal of Symbolic Logic 53 (1988), no. 2, pp. 434–442 (in French).

[3] Samuel R. Buss, *The Boolean formula value problem is in ALOGTIME*, in: Proceedings of the 19th Annual ACM Symposium on Theory of Computing, ACM Press, 1987, pp. 123–131.

[4] Patrick Cégielski, Serge Grigorieff, and Denis Richard, *La théorie élémentaire de la fonction de couplage de Cantor des entiers naturels est décidable*, Comptes Rendus de l'Académie des Sciences – Series I – Mathematics 331 (2000), no. 2, pp. 107–110 (in French).

[5] Patrick Cégielski and Denis Richard, *On arithmetical first-order theories allowing encoding and decoding of lists*, Theoretical Computer Science 222 (1999), no. 1–2, pp. 55–75.

[6] ——————, *Decidability of the theory of the natural integers with the Cantor pairing function and the successor*, Theoretical Computer Science 257 (2001), no. 1–2, pp. 51–77.

[7] Jeanne Ferrante and Charles W. Rackoff, *The computational complexity of logical theories*, Lecture Notes in Mathematics vol. 718, Springer-Verlag, 1979.

[8] Emil Jeřábek, *Answer to a question by Zuhair Al-Johar*, MathOverflow, 2018, `https://mathoverflow.net/a/302634`.

[9] Anatoliĭ I. Mal'cev, *On the elementary theories of locally free universal algebras*, Doklady Akademii Nauk SSSR 138 (1961), no. 5, pp. 1009–1012 (in Russian), English translation in: Soviet Mathematics – Doklady 2 (1961), no. 3, pp. 768–771.

[10] ——————, *Axiomatizable classes of locally free algebras of several types*, Sibirskiĭ Matematicheskiĭ Zhurnal 3 (1962), no. 5, pp. 729–743 (in Russian), English translation in: The Metamathematics of Algebraic Systems: Collected Papers: 1936–1967, Studies in Logic and the Foundations of Mathematics 66, North-Holland, 1971, pp. 262–281.

[11] Aleksei L. Semenov, *Logical theories of one-place functions on the set of natural numbers*, Izvestiya Akademii Nauk SSSR, Seriya Matematicheskaya 47 (1983), no. 3, pp. 623–658 (in Russian), English translation in: Mathematics of the USSR, Izvestiya 22 (1984), no. 3, pp. 587–618.

[12] Alfred Tarski, Andrzej Mostowski, and Rafael M. Robinson, *Undecidable theories*, North-Holland, Amsterdam, 1953.

[13] Richard L. Tenney, *Decidable pairing functions*, Ph.D. thesis, Cornell University, 1972.

[14] Robert L. Vaught, *Axiomatizability by a schema*, Journal of Symbolic Logic 32 (1967), no. 4, pp. 473–479.

[15] Albert Visser, *Pairs, sets and sequences in first-order theories*, Archive for Mathematical Logic 47 (2008), no. 4, pp. 299–326.