A second-order system for polytime reasoning based on Grädel's theorem

Stephen Cook and Antonina Kolokolova
University of Toronto
{sacook,kol}@cs.toronto.edu

January 10, 2002

Abstract

We introduce a second-order system V_1-Horn of bounded arithmetic formalizing polynomial-time reasoning, based on Grädel's [11] second-order Horn characterization of P. Our system has comprehension over P predicates (defined by Grädel's second-order Horn formulas) , and only finitely many function symbols. Other systems of polynomial-time reasoning either allow induction on NP predicates (such as Buss's S^1_2 or the second-order V^1_1), and hence are more powerful than our system (assuming the polynomial hierarchy does not collapse), or use Cobham's theorem to introduce function symbols for all polynomial-time functions (such as Cook's PV and Zambella's P-def). We prove that our system is equivalent to QPV and Zambella's P-def. Using our techniques, we also show that V_1-Horn is finitely axiomatizable, and, as a corollary, that the class of $\forall\Sigma^b_1$ consequences of S^1_2 is finitely axiomatizable as well, thus answering an open question.

1 Introduction

1.1 Bounded Arithmetic

The first theory that was explicitly designed in order for all proofs to be feasibly constructible (i.e., constructible in polynomial time) was the equational theory PV, proposed by Cook in 1975 [5]. There, Cobham's characterization of polynomial-time was used to construct polynomial-time functions. One motivation for PV was its close relation with Extended Frege proof systems for the propositional calculus: theorems of PV give rise to families of tautologies with polynomial-length proofs.

A major work establishing the relation between complexity theory and bounded arithmetic was the 1985 PhD thesis of S.Buss [2], where the first-order theories S^i_2 and various second-order theories were developed that characterize the levels of the polynomial-time hierarchy, PSPACE and EXPTIME. The most important of them is the first-order theory S^1_2, consisting of a set of 32 axioms and an induction-on-notation scheme over Σ^b_1 (NP) formulas. Buss proves that a function is Σ^b_1-definable in S^1_2 iff it is polynomial-time. He also shows that S^1_2 is $\forall\Sigma^b_1$ conservative over QPV (a quantified version of Cook's PV); however general conservativity of S^1_2 over QPV would imply the collapse of the polynomial hierarchy to P/poly [4, 16, 24]. Razborov and, at the same time, Takeuti [19, 22] introduce a general method (the RSUV isomorphism) for showing the equivalence between certain first-order and second-order theories, which can be used to show that the second-order theory V_1^1 is equivalent in power to S_2^1 [19]. Finally, in 1996 Zambella [24] introduced an elegant way of presenting a hierarchy of second-order theories equivalent to $\langle S_2^i \rangle$, as well as the second-order theory P-def, which includes function symbols for all polynomial-time functions and is equivalent to QPV.

1.2 Descriptive Complexity

While in bounded arithmetic we are interested in the proving power of theories, in descriptive complexity we are interested in expressive power of formulas. Instead of asking what class of theorems can be proven, we ask what properties (for example, graph properties) we can express using certain classes of formulas. This research goes back to Fagin's 1974 result [9] showing that a language is in NP iff it corresponds to the set of finite models of an existential second-order formula. Later Stockmeyer [21] extended this result, characterizing the polynomial hierarchy as the class of sets of finite models of all second-order formulas.

Finding an elegant descriptive-style characterization of P proved more illusive. One such characterization of P uses the first-order logic augmented with the successor relation and the least fixed-point operator [23, 13]. Later Leivant [17, 18] found a second-order characterization of P using the notion of "controlled computational formula", which is related to Horn formula. (The motivation for using Horn formulas comes from the existence of a simple polynomial-time algorithm for solving the satisfiability problem for propositional Horn formulas.) Finally Grädel [10, 11] found an elegant descriptive characterization of P using $SO\exists$-Horn (second-order existential Horn) formulas with successor.

1.3 Our Results

We present a second-order theory V_1-Horn of bounded arithmetic based on Grädel's theorem; our theory is intended to capture polynomial-time reasoning. We use the elegant syntax of Zambella's [24] second-order theories. Our main new feature is a comprehension axiom scheme for second-order existential Horn formulas, which by Grädel's theorem represent the polynomial-time predicates. Our main results are that V_1-Horn is finitely axiomatizable, from which it follows that the $\forall\Sigma_1^b$ consequences of S_2^1 are finitely axiomatizable, thus answering an open question (see [15], theorem 10.1.2). We also show that V^0 (Zambella's $\Sigma_0^b - comp$) is finitely axiomatizable. A major tool needed for our results is the construction in our theory of a Σ_1^B-Horn formula $\textsc{Run}_\Phi(R, \tilde{R})$ which encodes a run of the satisfiability algorithm on the Σ_1^B-Horn formula Φ (see Section 5).

Section 2 contains background information. We define our system V_1-Horn and other second-order theories in Section 3. In Section 4 we show that V_1-Horn proves the equivalence of each formula in several broad syntactic classes to a Σ_1^B-Horn formula. Section 5 contains the description of the main tool needed for later sections, namely representing the Horn satisfiability algorithm in V_1-Horn by a Σ_1^B-Horn formula. In Section 6 we construct a conservative extension V_1-Horn(FP) of V_1-Horn by introducing function symbols for polynomial-time functions, and show the equivalence of this and Zambella's P-def[24]. Finally, in Section 7 we demonstrate that both V^0 and V_1-Horn are finitely axiomatizable, and show that this implies that the $\forall\Sigma_1^b$ consequences of S_2^1 are finitely axiomatizable.

2 Second-order formulas and complexity classes

The prototype for the underlying language of V_1-Horn is the language of second-order bounded arithmetic introduced by Buss [2]. However our language is closer to the nicer second-order language introduced by Zambella [24], in that we eliminate the superscript terms t tagging second-order variables X^t and instead introduce a bounding function $|X|$.

Our language $\mathcal{L}_A^2$ has two sorts, called first-order and second-order. (The intention is that first-order objects are natural numbers and second-order objects are finite sets of natural numbers, or finite binary strings.) First-order variables are denoted by lower case letters $a, b, i, j, ..., x, y, z$, and second-order variables are denoted by upper-case letters $P, Q, ..., X, Y, Z$.

The first-order function and predicate symbols of $\mathcal{L}_A^2$ are the standard symbols $\{0, 1, +, \cdot; \leq, =\}$ of Peano Arithmetic. To these we add the unary upper-bound function symbol $|\ \ |$, which takes second-order objects to first-order objects, and the binary membership predicate symbol $\in$.

For every second-order variable X we form a first-order term $|X|$ called an *upper-bound term.* The first-order terms of $\mathcal{L}_A^2$ are built from 0, 1, first-order variables, and upper-bound terms using the function symbols + and $\cdot$. The only second-order terms are second-order variables.

The atomic formulas of $\mathcal{L}_A^2$ have one of the forms $s = t, s \leq t, t \in X$, where s and t are first-order terms and X is a second-order variable. We usually write $X(t)$ instead of $t \in X$. Formulas are built from atomic formulas using the propositional connectives $\wedge, \vee, \neg$, the first-order quantifiers $\forall x, \exists x$ and the second-order quantifiers $\forall X, \exists X$.

We use the usual abbreviations $s \neq t$ for $\neg s = t$ and $s < t$ for $s \leq t \wedge s \neq t$. Bounded first-order quantifiers get their usual meaning: $\forall x \leq t\phi$ stands for $\forall x(x \leq t \rightarrow \phi)$ and $\exists x \leq t\phi$ stands for $\exists x(x \leq t \wedge \phi)$. We also use bounded second order quantifiers: $\forall X \leq t\phi$ stands for $\forall X(|X| \leq t \rightarrow \phi)$ and $\exists X \leq t\phi$ stands for $\exists X(|X| \leq t \wedge \phi)$.

In the standard model for $\mathcal{L}_A^2$ first-order variables range over $\mathbb{N}$, and second-order variables range over finite subsets of $\mathbb{N}$. If X is the empty set, then $|X|$ is interpreted as 0, otherwise $|X|$ is interpreted as one more than the largest element of the finite set X. The symbols 0,1,+ $\cdot$, $\in$ get there usual interpretations.

In complexity theory a member of a language is often taken to be a binary string, but from our "second-order" point of view we take it to be a finite subset X of $\mathbb{N}$. To relate this to the string point of view we code a finite set X by the binary string X', where X' is the empty string if X is the empty set, and otherwise X' is the binary string $x_0x_1, ..., x_{n-1}$ of length $n = |X|$ such that $x_i = 1 \iff i \in X, 0 \leq i \leq n-1$. (Thus all nonempty string codes end in 1.) If L is a set of finite subsets of $\mathbb{N}$, then the corresponding set of strings is $L' = \{X' \mid X \in L\}$. If $\mathbf{C}$ is a standard complexity class such as $\mathbf{AC}^0$**,** $\mathbf{P}$ or $\mathbf{NP}$, then our second-order reinterpretation of $\mathbf{C}$ is $\{L \mid L' \in \mathbf{C}\}$. Since the complexity classes considered here are robust, this reinterpretation will come out the same for any reasonable string coding method.

The role of first-order objects in our theories is that of members of second-order objects, or equivalently as position indices for binary strings. Thus in determining the complexity of a set of natural numbers we code a natural number i using unary notation; that is as a string i' of 1's of length i.

Definition 2.1. *If $\phi(\bar{z}, \bar{Y})$ is a formula of $\mathcal{L}_A^2$ whose free variables are among $z_1, ..., z_k, Y_1, ..., Y_\ell$ then ϕ* represents *a $k+\ell$-ary relation R^ϕ as follows. If $a_1, ..., a_k$ are natural numbers and $B_1, ..., B_\ell$ are finite sets of natural numbers, then $\langle a_1, ..., a_k, B_1, ..., B_\ell\rangle$ satisfies R^ϕ iff $\phi(a_1, ..., a_k, B_1, ..., B_\ell)$ is true in the standard model.*

If $\mathbf{C}$ is a complexity class, then we make sense of the statement "R^ϕ is in $\mathbf{C}$" using the string encodings described above. In particular, a relation $R(x_1, ..., x_k, Y_1, ..., Y_m)$ is in $\mathbf{P}$ iff it is recognizable in time bounded by a polynomial in $(x_1, ..., x_k, |Y_1|, ..., |Y_m|)$.

We now define the classes Σ_i^B and Π_i^B of bounded second-order formulas. (A formula is *bounded* if all its quantifiers are bounded.) Σ_0^B and Π_0^B both denote the class of bounded formulas with no second-order quantifiers. We define inductively Σ_{i+1}^B as the least class of formulas containing Π_i^B and closed under disjunction, conjunction, and bounded existential second-order quantification . The class Π_{i+1}^B is defined dually.

The classes Σ_i^B and Π_i^B are the formulas in our (Zambella's) simplified language $\mathcal{L}_A^2$ which correspond to the classes $\Sigma_i^{1,b}$ and $\Pi_i^{1,b}$ in Buss's prototype second-order language [2, 15]. They are the second-order analogs of the first-order formula classes Σ_i^b and Π_i^b, where sharply-bounded quantifiers correspond to our bounded first-order quantifiers.

The formulas Σ_1^B represent precisely the $\mathbf{NP}$ relations, and more generally for $i > 1$ the Σ_i^B formulas represent the Σ_i^p relations in the polynomial hierarchy and Π_i^B represent the Π_i^p relations [2, 15]. The formulas Σ_0^B represent precisely the uniform $\mathbf{AC}^0$ relations, which are the same as the class $\mathbf{FO}$ (First Order) of descriptive complexity [1] (see Chapter 1 of [14]).

We now define the formulas corresponding to polynomial time.

Definition 2.2. *A formula ϕ of $\mathcal{L}_A^2$ is* Horn *with respect to the second-order variables $P_1, ..., P_k$ if ϕ is quantifier-free in conjunctive normal form and in every clause there is at most one positive literal of the form $P_i(t)$ (called the* head *of the clause) and no terms of the form $|P_i|$. (We do allow upper-bound terms $|X|$ and any number of positive literals $X(t)$, where X is not among $\{P_1, ..., P_k\}$.) A formula is Σ_1^B-Horn if it has the form*

$$\exists P_1 ... \exists P_k \forall x_1 \le t_1 ... \forall x_m \le t_m \phi \tag{1}$$

*where $k, m \ge 0$ and ϕ is Horn with respect to $P_1, ..., P_k$, and the bounding terms t_i do not involve $x_1, ..., x_m$. More generally a formula is Σ^B-*Horn *if it has the above form except that each second-order quantifier can be either $\exists$ or $\forall$. A formula is Π_1^b* Horn with respect to *$P_1, ..., P_k$ if it has the form (1) with the existential quantifiers omitted.*

Remark 2.3. *Note that our definition of Σ_1^B-Horn is somewhat different from the original Grädel's definition of second-order existential Horn formulas. Since our setting is that of bounded arithmetic rather than finite model theory, we bound all quantifiers (explicit bounds on first-order quantifiers give implicit bounds on second-order ones). We include both $+$ and $\times$ as interpreted functions (thus allowing pairing functions), while Grädel includes only successor. Grädel allows k-ary predicate symbols for each k, while we allow only unary predicate symbols (but we can simulate k-ary symbols using pairing functions).*

Notice that the second-order quantifiers in Σ_1^B-Horn and Σ^B-Horn formulas are not bounded. However, since no occurrence of $|P_i|$ is allowed, each such formula is equivalent in the standard model to one in which every quantifier $\exists P_i$ or $\forall P_i$ is bounded by a term t which is an upper bound on all terms u such that $P_i(u)$ occurs in the formula. On the other hand, if occurrences of $|P_i|$ were allowed, then an unbounded quantifier $\exists P_i$ can code an unbounded number quantifier $\exists |P_i|$ and hence undecidable relations would be representable.

It is often convenient to treat second-order objects as multi-dimensional arrays, instead of one-dimensional strings or sets. An easy way to do so is to use a pairing function $< \cdot, \cdot >$, defined by

$$< x, y >= (x+y)(x+y+1) + 2y \tag{2}$$

This function is a one-one map from $\mathbb{N} \times \mathbb{N}$ into $\mathbb{N}$, and it is represented by a term in our language. It is easily generalized to k-tuples by defining $\langle x_1, ..., x_k \rangle$ by the recursion

$$\langle x \rangle = x, \qquad \langle x_1, ..., x_{k+1} \rangle = \langle \langle x_1, ..., x_k \rangle, x_{k+1} \rangle \tag{3}$$

Thus, any finite set P can be treated as a set of k-tuples of variables; $P(x_1, ..., x_k)$ is defined to be $P(< x_1, ..., x_k >)$.

The theorem below is similar to part of Grädel's Theorem 5.2 [10] (see also Chapter 7 of [20]), which is stated in the context of descriptive complexity theory. There are technical differences: Grädel's language is more general in that it allows predicate symbols of arbitrary arity, but these can be simulated by the pairing function as just explained. On the other hand our language is more general in that it allows interpreted function symbols $+$ and $\cdot$ and terms $|Y_i|$, as well as number variables whose range goes up to any polynomial in the size of the inputs. However none of these generalizations takes us outside the polynomial-time relations.

Theorem 2.4. *A relation $R(z_1, ..., z_k, Y_1, ..., Y_m)$ is in* **P** *iff it is representable by a Σ_1^B-Horn formula Ψ. Further Ψ can be chosen with only one existentially quantified second-order variable, and only two universally quantified first-order variables.*

Example. (Parity(X)) This is a Σ_1^B-Horn formula which is true for strings X that contain an odd number of 1's. It encodes a dynamic-programming algorithm for computing parity of X: $P_{odd}(i)$ is true (and $P_{even}(i)$ is false) iff the prefix of X of length i contains an odd number of 1's.

$$\begin{aligned}
&\exists P_{even} \exists P_{odd} \forall i < |X| \\
&\qquad P_{even}(0) \wedge \neg P_{odd}(0) \wedge P_{odd}(|X|) \\
&\qquad \wedge(\neg P_{even}(i+1) \vee \neg P_{odd}(i+1)) \\
&\qquad \wedge(P_{even}(i) \wedge X(i) \to P_{odd}(i+1)) \wedge (P_{odd}(i) \wedge X(i) \to P_{even}(i+1)) \\
&\qquad \wedge(P_{even}(i) \wedge \neg X(i) \to P_{even}(i+1)) \wedge (P_{odd}(i) \wedge \neg X(i) \to P_{odd}(i+1))
\end{aligned}$$

Proof of theorem. For the if direction, let $\Psi(\bar{z}, \bar{Y})$ be a Σ_1^B-Horn formula which represents $R(\bar{z}, \bar{Y})$. Then Ψ has the form

$$\exists P_1 ... \exists P_r \forall x_1 \leq t_1 ... \forall x_s \leq t_s \phi(\bar{x}, \bar{P}, \bar{z}, \bar{Y}) \tag{4}$$

where ϕ is Horn with respect to $P_1, ..., P_r$. We outline a polynomial-time algorithm which, given numbers $a_1, ..., a_k$ (coded in unary) and finite sets $B_1, ..., B_m$ (coded by binary strings) determines whether $\Psi(\bar{a}, \bar{B})$ is true in the standard model. First note since $\bar{a}$ and $\bar{B}$ are given, each first-order term u in $\phi(\bar{x}, \bar{P}, \bar{a}, \bar{B})$ becomes a polynomial $u(x_1, ..., x_k)$, and the coefficients can be computed in polynomial-time. Each P_i can occur only in the context $P_i(u(\bar{x}))$ for some such term u, and the terms $t_1, ..., t_s$ bounding the x_i's evaluate to constants.

The algorithm proceeds by computing for each possible $\bar{x}$-value $\bar{b} = (b_1, ..., b_s), 0 \leq b_i \leq t_i$, a simplified form $\phi[\bar{b}]$ of the instance $\phi(\bar{b}, \bar{P}, \bar{a}, \bar{B})$ of ϕ. In this form all first-order terms and all atomic formulas not involving the P_i's are evaluated, and the result is a Horn formula $\phi[\bar{b}]$ all of whose atoms are in the list $P_i(0), ..., P_i(T), i = 1, ..., r$, where T is the largest possible argument of any P_i

in any instance. By taking the conjunction over all $\bar{b}$ of these instances, we obtain a propositional Horn formula PROP$[\phi, \bar{a}, \bar{B}]$. It is not hard to see that $\Psi(\bar{a}, \bar{B})$ is true in the standard model iff PROP$[\phi, \bar{a}, \bar{B}]$ is satisfiable.

Finally, there is a standard polynomial-time algorithm to test satisfiability of a given propositional Horn formula Φ. Namely, initialize a truth assignment τ to set all atoms to false. Now repeatedly, for each clause C in Φ not satisfied by the current τ, either C has no positive occurrence of an atom P, in which case Φ is unsatisfiable, or C has a unique positive occurrence of some atom P, in which case flip the value of τ on P from false to true.

The proof of the only-if direction resembles the proof of Cook's theorem that SAT is **NP**-complete, and of Fagin's theorem of finite model theory that second-order existential formulas capture **NP**. Let M be a deterministic Turing machine that recognizes a relation $R(x_1, ..., x_k, Y_1, ..., Y_m)$ within time n^ℓ, where $n = x_1 + ... + x_k + |Y_1| + ... + |Y_m|$ is the length of the input. The entire computation of M on this input can be represented by a two dimensional array $P(i, j)$ with $t(n)$ rows and columns, for some polynomial t, where the i-th row specifies the tape configuration at time i. (P can be represented by a one-dimensional array using a pairing function, as explained above.) Thus $R(\bar{x}, \bar{Y})$ is represented by the Σ_1^B-Horn formula

$$\exists P \exists \tilde{P} \forall i \leq t(n) \forall j \leq t(n) \phi(P, \tilde{P}, i, j, \bar{x}, \bar{Y}) \tag{5}$$

Here the variable $\tilde{P}$ is forced to be $\neg P$ in the same way that P_{even} and P_{odd} are forced to be complementary in the parity example above. The formula $\phi(P, \tilde{P}, i, j, \bar{x}, \bar{Y})$ is Horn with respect to P and $\tilde{P}$, and each clause specifies a local condition on the computation. These conditions are (1) the first row of P codes the initial tape configuration for the inputs $\bar{x}, \bar{Y}$, (2) for $i < t(n)$ the $i+1$-st row represents the i-th row after one step, and (3) the final state is accepting. To make (2) easier to specify, it is convenient to represent the state at time i at the beginning of row i by a string of fixed length, and after the code for the symbol stored at each tape position there is a bit specifying whether that square is currently scanned by the Turing machine head. In this way rows i and $i+1$ will be identical except for the state codes at the beginning and the bits coding the old and new tape squares scanned.

To see that each clause can be designed to meet the Horn condition of at most one positive occurrence among the atoms of the form $P(u), \tilde{P}(u)$, we include the clause $(\neg P(i, j) \vee \neg \tilde{P}(i, j)$. Then every bit in row 0 is specified using a clause with a positive literal of one of the forms $P(0, u)$ or $\tilde{P}(0, u)$, possibly together with other literals involving input variables. For example, if 15 bits are reserved at the beginning of each row to specify the state, and 3 bits code each tape square, then one of the clauses might be $(5 \leq j \wedge j \leq 5 + x_1 \to P(0, 3 \cdot j + 1))$. In general every bit in row $i+1$ is specified conditional on a fixed number of bits in row i. A clause is included for each possible state of these conditional bits, and the conditions are specified using $\neg P$ and $\neg \tilde{P}$ as appropriate. In this way at least one of $P(i, j), \tilde{P}(i, j)$ must be true for each (i, j) (and hence exactly one). Note however that if M were nondeterministic, then row $i+1$ would have more than one possible value, and some clauses would require more than one positive literal so the formula would not be Horn.

To meet the "further" condition stated in the theorem, the two arrays P and $\tilde{P}$ can be combined into one array $Q(i, j, k)$, where $k = 0$ for P and $k = 1$ for $\tilde{P}$. □

Note that above proof also shows that every **NP**- relation can be represented by a Σ_1^B formula of the form (5), except that ϕ is not Horn.

Example. (3Color(n, E)) This is a Σ_1^B formula asserting that the graph with edge relation E on

Robinson's Theory Q axioms	B1	$x+1 \neq 0$	B2	$x+1 = y+1 \rightarrow x = y$
	B3	$x+0 = x$	B4	$x+(y+1) = (x+y)+1$
	B5	$x \cdot 0 = 0$	B6	$x \cdot (y+1) = (x \cdot y) + x$
Axioms for $\leq$	B7	$0 \leq x$	B8	$x \leq x+y$
	B9	$x \leq y \wedge y \leq z \rightarrow x \leq z$	B10	$(x \leq y \wedge y \leq x) \rightarrow x = y$
	B11	$x \leq y \vee y \leq x$	B12	$x \leq y \leftrightarrow x < y+1$
Predecessor	B13	$x \neq 0 \rightarrow \exists y(y+1 = x)$		
Upper bound	L1	$X(y) \rightarrow y < \lvert X\rvert$	L2	$y+1 = \lvert X\rvert \rightarrow X(y)$

Table 1: The 2-*BASIC* Axioms

nodes $\{0, 1, ..., n-1\}$ is three-colorable. We write $E(x, y)$ like a binary relation, although it can be coded as a unary relation using the pairing function as explained above. The three colors are P, Q, and R.

$$\exists P \exists Q \exists R \forall x < n \forall y < n(P(x) \vee Q(x) \vee R(x))$$
$$\wedge(\neg E(x,y) \vee \neg P(x) \vee \neg P(y)) \wedge (\neg E(x,y) \vee \neg Q(x) \vee \neg Q(y)) \wedge (\neg E(x,y) \vee \neg R(x) \vee \neg R(y))$$

This formula is Σ_1^B-Horn except for the first clause. Since graph 3-colorability is **NP**-compete, it cannot be represented by a Σ_1^B-Horn formula unless $\mathbf{P} = \mathbf{NP}$. This example illustrates why we cannot allow bounded first-order existential quantifiers after the universal quantifiers in Σ_1^B-Horn formulas, since the first clause could be replaced by $\exists i < 3P(i, x)$ where now $P(0, x), P(1, x), P(2, x)$ represent the three colors.

3 V_1-Horn and other second-order theories

Our second-order theories use the language $\mathcal{L}_A^2$ described in the previous section. They all share the set 2-*BASIC* of axioms in Table 1, which are similar to the axioms for Zambella's theory Θ [24] and form the second-order analog of Buss's first-order axioms *BASIC* [2]. The set 2-*BASIC* consists essentially of the axioms for Robinson's system Q, together with axioms for $\leq$, and two axioms defining the upper-bound terms $|X|$.

In addition to 2-*BASIC*, each system needs a comprehension scheme for some set *FORM* of formulas.

$$FORM - \text{COMP} : \exists X \leq y \forall z < y(X(z) \leftrightarrow \Phi(z)) \tag{6}$$

Here, Φ is any formula in the set *FORM* with no free occurrence of X.

We denote by V^i the theory axiomatized by 2-*BASIC* and Σ_i^B-COMP. For $i \geq 0$ V^i is essentially the same as Zambella's $\Sigma_i^p - comp$ [24]. For $i \geq 1$ V^i is essentially the same as V_1^i [15]. (The latter restricts comprehension to $\Sigma_0^{1,b}$ formulas, but allows induction on $\Sigma_i^{1,b}$ formulas. However Theorem 1 of Buss [3] shows that V_1^i proves the $\Sigma_i^{1,b}$ comprehension axioms.) Thus for $i \geq 1$ V^i is a second-order version of S_2^i. In particular, the Σ_1^B-definable functions in V^1 are precisely the polynomial-time functions [6]. The Σ_1^B-definable functions in V^0 are the uniform $\mathbf{AC}^0$ functions [6] (called *rudimentary functions* in [24]). The first-order analog of V^0 is S_2^0 with a comprehension scheme for sharply-bounded formulas.

Definition 3.1. *V_1-Horn is the theory axiomatized by 2-BASIC and Σ_1^B-Horn-COMP.*

Although 2-$BASIC$ does not include an explicit induction axiom, L2 asserts that a nonempty set has a largest element. This can be turned into a least number principle, from which induction follows.

Lemma 3.2. *The least number principle is a theorem of V_1-Horn, and of $V^i, i \geq 0$.*

$$\textit{LNP: } 0 < |X| \rightarrow \exists x < |X|(X(x) \wedge \forall y < x \neg X(y))$$

Proof. By the comprehension schema there is a set Y such that $|Y| \leq |X|$ and for all $z < |X|$

$$Y(z) \leftrightarrow \forall i < |X|(X(i) \rightarrow z < i)$$

Thus the set Y consists of those elements smaller than every element in X. We claim that $|Y|$ satisfies the LNP for X; that is (i) $|Y| < |X|$, (ii) $X(|Y|)$ and (iii) $\forall y < |Y| \neg X(y)$. First suppose that Y is empty. Then $|Y| = 0$ by B13 and L2. By assumption $0 < |X|$, so (i) holds in this case. Also $X(0)$, since otherwise $Y(0)$ by B7 and the definition of Y, so (ii) holds. Since $\neg y < 0$ by B7 and B10 we conclude (iii) holds vacuously.

Now suppose $Y(y)$ for some y. Then $y < |Y|$ by L1, so $|Y| \neq 0$ so by B13 $|Y| = z + 1$ for some z and hence $Y(z)$ by L2. Then $\neg Y(z+1)$ by L1. Thus $X(z+1)$ by B11, B12 and the definition of Y, so (ii) holds. Also $\neg X(z)$, so (i) holds. Finally (iii) holds by the definition of Y and B10. □

Lemma 3.3. *Induction on length of a string is a theorem of V_1-Horn, and of $V^i, \geq 0$.*

$$\textit{IND: } (X(0) \wedge \forall y < z(X(y) \rightarrow X(y+1))) \rightarrow X(z)$$

Proof. We show $\neg$IND $\rightarrow$ $\neg$LNP. By $\neg$IND we have $X(0) \wedge \forall y < z(X(y) \rightarrow X(y+1)$, and $\neg X(z)$. By the comprehension schema there is a set Y such that $\forall y < z+1(Y(y) \leftrightarrow \neg X(y)$. Then $Y(z)$, so $0 < |Y|$. By LNP Y has a least element y_0. Then $y_0 \neq 0$ because $X(0)$, so $y_0 = x_0 + 1$ for some x_0, by B13. But then we must have $X(x_0)$ and $\neg X(x_0 + 1)$, which contradicts our assumption. □

This is easy to generalize to allow induction with an arbitrary k as a basis, not just $k = 0$.

If follows from the above Lemma that each of the theories that we have presented proves an induction axiom for each formula in its comprehension scheme. In particular, for V_1-Horn we have

Corollary 3.4. *V_1-Horn proves the Σ_1^B-Horn Induction axioms.*

$$\Sigma_1^B\textit{-Horn-IND}: (\Phi(0) \wedge \forall y < z(\Phi(y) \rightarrow \Phi(y+1))) \rightarrow \Phi(z)$$

where Φ is any Σ_1^B-Horn formula.

Standard arguments show that induction on open formulas using axioms B1 to B13 is enough to prove simple algebraic properties of $+$ and $\cdot$ such as commutativity, associativity, distributive laws, and cancellation laws involving $+$, $\cdot$, and $\leq$. Hence all of our theories prove these properties, and in the sequel we take them for granted. These simple properties suffice to prove that the tupling function defined in (2) and (3) is one-one, so these theories all prove

$$\langle x_1, ..., x_k \rangle = \langle x'_1, ..., x'_k \rangle \rightarrow (x_1 = x'_1 \wedge ... \wedge x_k = x'_k) \tag{7}$$

Lemma 3.5 (k-ary Comprehension)**.** *If $\Phi(x_1, ..., x_k)$ is a Σ_1^B-Horn formula with no free occurrence of Y, then V_1-Horn proves the k-ary comprehension formula*

$$\exists Y \leq \langle b_1, ..., b_k \rangle \forall x_1 < b_1 ... \forall x_k < b_k (Y(x_1, ..., x_k) \leftrightarrow \Phi(x_1, ..., x_k)) \tag{8}$$

Proof. Let $\Psi(i)$ be the formula

$$\forall x_1 < b_1 ... \forall x_k < b_k (i = \langle x_1, ..., x_k \rangle \rightarrow \Phi(x_1, ..., x_k))$$

Then the prenex form of $\Psi(i)$ is Σ_1^B-Horn, so by the comprehension scheme V_1-Horn proves the existence of a set Y such that $|Y| \leq t(b_1, ..., b_k)$ and $\forall i < t(b_1, ..., b_k)(Y(i) \leftrightarrow \Psi(i))$. Thus V_1-Horn proves $\forall x < b_1 ... \forall x_k < b_k (Y(x_1, ..., x_k) \leftrightarrow \Phi(x_1, ..., x_k))$, using the fact (7) that the tupling function is one-one. □

4 Formulas provably equivalent to Σ_1^B-Horn

Our goal now is to show that every Σ_0^B formula and every Σ_i^B-Horn formula, $i \in \mathbb{N}$, is provably equivalent in V_1-Horn to a Σ_1^B-Horn formula, and hence can be used in the comprehension and induction schemes. Later, we also show that the class of formulas provably equivalent to Σ_1^B-Horn is closed under $\neg, \wedge, \vee$ and bounded first-order quantification (see 5.3). We start with a simple observation.

Lemma 4.1. *If Φ_1 and Φ_2 are Σ_1^B-Horn formulas, then $\Phi_1 \wedge \Phi_2$ is logically equivalent to a Σ_1^B-Horn formula.*

Proof. Take a suitable prenex form of $\Phi_1 \wedge \Phi_2$. □

4.1 Simulating first-order bounded existential quantification

A major inconvenience of Σ_1^B-Horn formulas is lack of first-order existential quantifiers. In general we cannot allow such quantifiers without increasing the apparent expressive power of the formulas, as pointed out in the 3-colorability example. However, it is possible to introduce bounded existential quantifiers in some contexts.

Notation. If P is a second-order variable, then $\tilde{P}$ denotes a second-order variable whose intended interpretation is $\neg P$. □

We now introduce the Horn formulas $\textsc{Search}_k$, which are Π_1^b Horn with respect to all of their second-order variables and which will allow a Σ_1^B-Horn formula to represent $\exists z < b X(\bar{y}, z)$. $\textsc{Search}_k(\bar{b}, b, S, \tilde{S}, X, \tilde{X})$ asserts that $S(\bar{y}, i)$ holds iff $X(\bar{y}, z)$ holds for some $z < i$, where $\bar{b}$ stands for $b_1, ..., b_k$, and $\bar{y}$ stands for $y_1, ..., y_k$. We use $\bar{y} < \bar{b}$ for $y_1 < b_1 \wedge ... \wedge y_k < b_k$.

Definition 4.2. *For each $k \geq 1$ $\textsc{Search}_k(\bar{b}, b, S, \tilde{S}, X, \tilde{X})$ is the Π_1^b Horn formula*

$$\begin{aligned}
\forall \bar{y} < \bar{b} \forall i < b (&\neg S(\bar{y}, 0) \wedge \tilde{S}(\bar{y}, 0)) \\
&\wedge (\neg S(\bar{y}, i+1) \vee \neg \tilde{S}(\bar{y}, i+1)) \\
&\wedge (S(\bar{y}, i) \rightarrow S(\bar{y}, i+1)) \\
&\wedge (X(\bar{y}, i) \rightarrow S(\bar{y}, i+1)) \\
&\wedge (\tilde{S}(\bar{y}, i) \wedge \tilde{X}(\bar{y}, i) \rightarrow \tilde{S}(\bar{y}, i+1))
\end{aligned}$$

Lemma 4.3. *V_1-Horn proves the following:*

$$(i)\ \ \forall z < b(X(\bar{y}, z) \leftrightarrow \neg\tilde{X}(\bar{y}, z)) \wedge \bar{y} < \bar{b} \ \rightarrow\ \exists S \exists \tilde{S}\ \textsc{Search}_k(\bar{b}, b, S, \tilde{S}, X, \tilde{X})$$

$$(ii)\ \forall z < b(X(\bar{y}, z) \leftrightarrow \neg\tilde{X}(\bar{y}, z)) \wedge \textsc{Search}_k(\bar{b}, b, S, \tilde{S}, X, \tilde{X}) \wedge \bar{y} < \bar{b}$$
$$\rightarrow (S(\bar{y}, b) \leftrightarrow \exists z < bX(\bar{y}, z)) \wedge (\tilde{S}(\bar{y}, b) \leftrightarrow \forall z < b\tilde{X}(\bar{y}, z))$$

Proof. First we prove (i). Arguing in V_1-Horn, there are two cases. If $\forall z < b\tilde{X}(\bar{y}, b)$ then use $k+1$-ary comprehension (Lemma 3.5) to define $S(\bar{y}, z)$ false and $\tilde{S}(\bar{y}, z)$ true, for all $z < b$. The clauses in the definition of $\textsc{Search}_k$ are clearly satisfied in this case. Otherwise, by the LNP there is a least number $z_0 < b$ such that $X(\bar{y}, x_0)$. Use $k+1$-ary comprehension to define $S(\bar{y}, z)$ false for $z \leq z_0$ and true for $z_0 < z < b$, and define $\tilde{S}(\bar{y}, z) \leftrightarrow \neg S(\bar{b}, z)$. Again $\textsc{Search}_k(\bar{b}, b, S, \tilde{S}, X, \tilde{X})$ holds.

To prove (ii) we use the same two cases as for (i). If $\forall z < b\tilde{X}(\bar{y}, b)$ we use the definition of $\textsc{Search}_k$ to show by induction on z that $S(\bar{y}, z)$ is false and $\tilde{S}(\bar{y}, z)$ is true for $z \leq b$, so (ii) holds in this case. For the second case we know from above what S and $\tilde{S}$ must be, and we again prove our claim by induction on z. Again (ii) follows. □

4.2 The Σ_0^B formulas are provably equivalent to Σ_1^B-Horn

Consider a Σ_0^B formula $Q_1 y_1 < b_1 ... Q_k y_k < b_k \phi(\bar{y})$, where each Q_i is either $\forall$ or $\exists$. The proof of the following Lemma shows how to conjoin copies of $\textsc{Search}(...)$ to define arrays $S_0, ..., S_k$ such that $S_i(y_1, ..., y_{k-i}) \leftrightarrow Q_{k-i+1} y_{k-i+1} < b_{k-i+1} \phi(\bar{y})$. These are used to form an equivalent Σ_1^B-Horn formula.

Lemma 4.4. *Let $\psi(\bar{y})$ be a Σ_0^B formula which may have other free variables besides $\bar{y}$ but does not involve any of the variables $S, \tilde{S}, \bar{W}$. Then there is a formula $\psi^*(\bar{b}, S, \tilde{S}, \bar{W})$ not involving $\bar{y}$ but which may have other variables of ψ not indicated and which is Π_1^b Horn with respect to $S, \tilde{S}, \bar{W}$ such that V_1-Horn proves the following:*

$$(i)\ \ \exists S \exists \tilde{S} \exists \bar{W} \psi^*(\bar{b}, S, \tilde{S}, \bar{W})$$
$$(ii)\ \psi^*(\bar{b}, S, \tilde{S}, \bar{W}) \rightarrow \forall \bar{y} < \bar{b}[(S(\bar{y}) \leftrightarrow \psi(\bar{y})) \wedge (\tilde{S}(\bar{y}) \leftrightarrow \neg\psi(\bar{y}))]$$

Proof. We may assume that ψ is in prenex form, and proceed by induction on the number of quantifiers. For the base case ψ is quantifier-free, and we take $\psi^*(\bar{b}, S, \tilde{S})$ to be equivalent to

$$\forall \bar{y} < \bar{b}[(S(\bar{y}) \leftrightarrow \psi(\bar{y})) \wedge (\tilde{S}(\bar{y}) \leftrightarrow \neg\psi(\bar{y}))]$$

The formula in brackets can be written in conjunctive normal form, in which case $\psi^*(\bar{b}, S, \tilde{S})$ is Π_1^b Horn with respect to S and $\tilde{S}$ and obviously satisfies (ii). Also (i) is easily proved by defining S and $\tilde{S}$ using Σ_1^B-Horn comprehension.

For the induction step, assume that $\psi(\bar{y})$ is $\exists z < t\phi(\bar{y}, z)$, where t is a term not involving z. By the induction hypothesis applied to ϕ there is a formula $\phi^*(\bar{b}, b, S_1, \tilde{S}_1, \bar{W})$ not involving $\bar{y}, z$ which is Π_1^b Horn with respect to $S_1, \tilde{S}_1, \bar{W}$ which satisfies (i) and (ii) (with ϕ, ϕ^*, S_1 for ψ, ψ^*, S). In fact the induction hypothesis (ii) states

$$\phi^*(\bar{b}, b, S_1, \tilde{S}_1, \bar{W}) \rightarrow \forall \bar{y} < \bar{b} \forall z < b(S_1(\bar{y}, z) \leftrightarrow \phi(\bar{y}, z)) \wedge (\tilde{S}_1(\bar{y}, z) \leftrightarrow \neg\phi(\bar{y}, z))$$

We define $\psi^*(\bar{b}, S, \tilde{S}, S_1, \tilde{S}_1, \bar{W})$ to be the prenex form of

$$\phi^*(\bar{b}, t, S_1, \tilde{S}_1, \bar{W}) \wedge \textsc{Search}_k(\bar{b}, t, S, \tilde{S}, S_1, \tilde{S}_1) \tag{9}$$

Note that this is Π_1^b Horn with respect to the displayed second-order variables. By the induction hypothesis (i) there exists $S_1, \tilde{S}_1, \bar{W}$ satisfying ϕ^*. By the induction hypothesis (ii) we have $S_1 \leftrightarrow \neg\tilde{S}_1$. Hence by (i) of Lemma 4.3 we know $S, \tilde{S}$ exist satisfying (i) in the present Lemma for ψ^* as defined above.

To prove (ii), assume $\bar{y} < \bar{b}$ and $\psi^*(\bar{b}, S, \tilde{S}, S_1, \tilde{S}_1, \bar{W})$. By the induction hypothesis (ii) for ϕ^* and (ii) of Lemma 4.3 we have $S(\bar{y}, t) \leftrightarrow \exists z < t\phi(\bar{y}, z)$ and $\tilde{S}(\bar{y}, t) \leftrightarrow \forall z < b\neg\phi(\bar{y}, z)$, as required.

For the induction step in case $\psi(\bar{y})$ is $\forall z < t\phi(\bar{y}, z)$ we simply modify the arguments of $\textsc{Search}_k$ in (9) by interchanging S with $\tilde{S}$ and S_1 with $\tilde{S}_1$. $\square$

Corollary 4.5. *Every Σ_0^B formula is provably equivalent in V_1-Horn to a Σ_1^B-Horn formula.*

Proof. Let ψ be a Σ_0^B formula not involving y and let $\psi^*(b, S, \tilde{S}, \bar{W})$ result from applying the above Lemma to $\psi(y)$. Then $\psi(y) \leftrightarrow \psi(0)$ so V_1-Horn proves

$$\psi(y) \leftrightarrow \exists S \exists \tilde{S} \exists \bar{W}(\psi^*(1, S, \tilde{S}, \bar{W}) \wedge S(0))$$

The right hand side is easily equivalent to a Σ_1^B-Horn formula. $\square$

Thus V_1-Horn proves the induction and comprehension schemes for Σ_0^B formulas, and hence it is an extension of V^0.

4.3 Collapse of V-Horn to V_1-Horn

Grädel [10] showed that it is possible to represent a $SO\exists$-Horn formula preceded by alternating SO quantifiers by a $SO\exists$-Horn formula, which implies the collapse of SO-Horn hierarchy to $SO\exists$-Horn. Here we formalize Grädel's proof in V_1-Horn.

First we show that V_1-Horn proves a version of the replacement scheme.

Notation. We use $P^{[b]}$ to denote the "b-th row" when P is being used as a 2-dimensional array. If $\phi(P)$ is a formula with no occurrence of $|P|$, then $\phi(P^{[b]})$ is obtained from $\phi(P)$ by replacing every atomic formula $P(t)$ by $P(b, t)$ (i.e. $P(\langle b, t\rangle)$: see (2)).

Lemma 4.6 (Replacement)**.** *If $\phi(y, \bar{P})$ is a Π_1^b Horn formula with respect to $\bar{P}$ and t is a term not involving y, then V_1-Horn proves*

$$\forall y < t \exists \bar{P} \phi(y, \bar{P}) \leftrightarrow \exists \bar{P} \forall y < t \phi(y, \bar{P}^{[y]})$$

where $\bar{P}^{[y]}$ is $P_1^{[y]}, ..., P_k^{[y]}$. Further the RHS is a Σ_1^B-Horn formula.

Proof. The last statement is immediate from the definition of Σ_1^B-Horn formula. To prove the first statement we move the quantifier $\forall y < t$ past each $\exists P_i$ in turn, using the following lemma. $\square$

Lemma 4.7. *If V_1-Horn proves that $\exists P \forall y < b\Phi(y, P)$ is equivalent to some Σ_1^B-Horn then V_1-Horn proves*

$$\forall y < b \exists P \Phi(y, P) \leftrightarrow \exists P \forall y < b \Phi(y, P^{[y]})$$

Proof. To prove the right-to-left implication, assume that P satisfies the existential quantifier on the right and suppose $y < b$. Use the V_1-Horn comprehension axiom to define P' such that

$$\forall i < b(P'(i) \leftrightarrow P(y,i))$$

Then P' satisfies the existential quantifier on the left.

To prove the left-to-right direction define

$$\Psi(z) \equiv \exists P \forall y < z \Phi(y, P^{[y]})$$

Then by assumption $\Psi(z)$ is equivalent to a SO$\exists$-Horn formula, so we may use the IND scheme (Corollary 3.4) to conclude $\Psi(b)$. It suffices to prove that the LHS $\forall y < b \exists P \Phi(y,P)$ implies the basis and induction steps. The basis is trivial, since when $b = 0$ $\Psi(0)$ is vacuously true.

For the induction step, by the induction hypothesis $\Psi(z)$ we may assume $z < b$ and P satisfies $\forall y < z\Phi(y, P^{[y]})$. Setting $y = z$ in the LHS we have Q such that $\Phi(z, Q)$. Now we use binary comprehension (Lemma 3.5) to define $P'(y,i)$ by

$$P'(y,i) \leftrightarrow \begin{cases} P(y,i) & \text{if } y < z \\ Q(i) & \text{if } y = z \end{cases}$$

Then we conclude in V_1-Horn the formula $\forall y < z + 1\Phi(y, P'^{[y]})$, and hence $\Psi(z + 1)$. □

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

Theorem 4.8. *Every SO Horn formula is provably equivalent in V_1-Horn to a SO$\exists$-Horn formula.*

This follows from the Replacement Lemma and the following Lemma.

Lemma 4.9. *If $\phi(P, \bar{Q})$ is Π_1^b Horn with respect to $P, \bar{Q}$ then V_1-Horn proves*

$$\forall P \exists \bar{Q} \phi(P, \bar{Q}) \leftrightarrow \forall y \le u \exists \bar{Q} \phi'(y, \bar{Q})$$

where if $P(t_1), ..., P(t_k)$ is a list of all occurrences of P in ϕ, then u is the term $t_1 + ... + t_k + 1$, and $\phi'(y, \bar{Q})$ is obtained from $\phi(P, \bar{Q})$ by replacing each $P(t_i)$ by $t_i \neq y$.

Proof. First note that V_1-Horn proves $t_i < u$, for $i = 1, ..., k$. To prove the left-to-right direction, for each y simply use comprehension to define P by the condition

$$\forall i \le u(P(i) \leftrightarrow i \neq y)$$

The proof of the converse is more complicated. Given P we use Σ_0^B comprehension to define the sets $\bar{Q}$ in terms of P and the $\bar{Q}$ from the RHS. There are two cases. The easy case is that $\forall z < uP(z)$ holds. Then take $y = u$, and the $\bar{Q}$ which satisfy the RHS will also satisfy the LHS, since $t_i \neq y$ for each i.

Now suppose $\exists z < u \neg P(z)$. By the Replacement Lemma applied to the RHS there are $\bar{Q}'$ satisfying $\forall y \le u \phi'(y, \bar{Q}'^{[y]})$. For each $Q_j \in \bar{Q}$ use Σ_0^B comprehension to define Q_j by the condition

$$\forall z < u_j(Q_j(z) \leftrightarrow \forall y < u(P(y) \vee Q_j'^{[y]}(z)))$$

where u_j is an upper bound on all terms v such that $Q_j(v)$ occurs in ϕ.

It remains to argue in V_1-Horn that this definition of $\bar{Q}$ satisfies $\phi(P, \bar{Q})$. We argue the contrapositive: If $\neg\phi(P, \bar{Q})$ then $\neg\phi'(y, \bar{Q}'^{[y]})$ for some y. Recall that ϕ begins with a string of bounded universal quantifiers $\forall \bar{x} \leq \bar{w}$, followed by a quantifier-free formula ψ which is Horn with respect to $P, \bar{Q}$. Fix values for the variables $\bar{x}$ which cause some clause $C(\bar{x}, P, \bar{Q})$ in ψ to be false. We will show that the corresponding clause $C'(\bar{x}, y, \bar{Q}'^{[y]})$ in ϕ' is false for a suitable choice of y. If the head of C is $P(t_i)$, then take $y = t_i$. If the head of C is $Q_j(v)$, then choose $y \leq u$ satisfying $(\neg P(y) \wedge \neg Q_j'^{[y]}(v))$. Such a y must exist because $\neg Q_j(v)$. Otherwise choose any $y \leq u$. In each case it is easy to see that $C'(\bar{x}, y, \bar{Q}'^{[y]})$ is false. □

5 Encoding the Horn SAT algorithm by a Σ_1^B-Horn formula

Here we show that a run of the Horn satisfiability algorithm described in the proof of Theorem 2.4 can be represented by a Σ_1^B-Horn formula RUN. This result is needed for sections 6 and 7. A simple corollary is that the negation of a Σ_1^B-Horn formula is provably equivalent to a Σ_1^B-Horn formula. In other words, V_1-Horn proves that P is closed under complementation.

Theorem 5.1. *Let Φ be a Σ_1^B-Horn formula which does not involve R or $\tilde{R}$. Then there is a formula $\text{RUN}_\Phi(R, \tilde{R})$ whose free variables include those of Φ in which the only atomic subformulas involving R and $\tilde{R}$ are $R(0)$ and $\tilde{R}(0)$ and such that $\exists R \exists \tilde{R} \text{RUN}_\Phi(R, \tilde{R})$ is a Σ_1^B-Horn formula and V_1-Horn proves the following:*

$$(i)\ \ \exists R \exists \tilde{R}\ \text{RUN}_\Phi(R, \tilde{R})$$

$$(ii)\ \ \text{RUN}_\Phi(R, \tilde{R}) \to [(R(0) \leftrightarrow \Phi) \wedge (\tilde{R}(0) \leftrightarrow \neg\Phi)]$$

Corollary 5.2. *If Φ is Σ_1^B-Horn, then $\neg\Phi$ is provably equivalent in V_1-Horn to a Σ_1^B-Horn formula NEG_Φ.*

Proof. We may take NEG_Φ to be $\text{RUN}_\Phi(\bot, \top)$; that is $\text{RUN}_\Phi(R, \tilde{R})$ with each occurrence of the formula $R(0)$ replaced by $\bot$ (FALSE) and each occurrence of the formula $\tilde{R}(0)$ replaced by $\top$ (TRUE). □

Corollary 5.3. *The class of formulas provably equivalent in V_1-Horn to a Σ_1^B-Horn formula is closed under $\neg$, $\wedge$, $\vee$, and bounded first-order quantification.*

Proof. The preceding corollary handles the case of $\neg$, Lemma 4.1 handles the case of $\wedge$, and the Replacement Lemma 4.6 handles the case of $\forall y < t$. The other cases follow by DeMorgan's laws. □

Theorem 5.1 can be generalized to the case in which arrays $R(\bar{y})$ and $\tilde{R}(\bar{y})$ code values of $\Phi(\bar{y})$ and $\neg\Phi(\bar{y})$.

Corollary 5.4. *Let $\Phi(\bar{y})$ be a Σ_1^B-Horn formula which does not involve R or $\tilde{R}$. Then there is a formula $\text{RUN}_{\Phi(\bar{y})}(\bar{b}, R, \tilde{R})$ which does not have $\bar{y}$ free but whose free variables include any other free variables of Φ such that $\exists R \exists \tilde{R} \text{RUN}_{\Phi(\bar{y})}(R, \tilde{R})$ is a Σ_1^B-Horn formula and V_1-Horn proves the following:*

$$(i)\ \ \exists R \exists \tilde{R}\ \text{RUN}_{\Phi(\bar{y})}(\bar{b}, R, \tilde{R})$$

$$(ii)\ \ \text{RUN}_{\Phi(\bar{y})}(\bar{b}, R, \tilde{R}) \to \forall \bar{y} < \bar{b}[(R(\bar{y}) \leftrightarrow \Phi(\bar{y})) \wedge (\tilde{R}(\bar{y}) \leftrightarrow \neg\Phi(\bar{y}))]$$

Proof. We take $\textsc{Run}_{\Phi(\bar{y})}$ such that V_1-Horn proves

$$\textsc{Run}_{\Phi(\bar{y})}(\bar{b}, R, \tilde{R}) \leftrightarrow \forall \bar{y} < \bar{b} \exists R' \exists \tilde{R}' [\textsc{Run}_{\Phi}(R', \tilde{R}') \wedge (R(\bar{y}) \leftrightarrow R'(0)) \wedge (\tilde{R}(\bar{y}) \leftrightarrow \tilde{R}'(0))]$$

We may take $\textsc{Run}_{\Phi(\bar{y})}$ to be Σ_1^B-Horn by placing the subformula enclosed in [...] above by a suitable prenex form and applying Corollary 5.3. To prove (i) we use Σ_1^B-Horn comprehension to define $R(\bar{y})$ satisfying $R(\bar{y}) \leftrightarrow \Phi(\bar{y})$ and use Σ_1^B-Horn comprehension together with corollary 5.2 to define $R(\bar{y}) \leftrightarrow \neg\Phi(\bar{y})$ and then apply (i) and (ii) of Theorem 5.1 to R' and $\tilde{R}'$. To prove (ii) we use (ii) in Theorem 5.1. □

We begin the proof of Theorem 5.1 by observing that one existential quantifier is enough in a Σ_1^B-Horn formula. (Recall the notation $P^{[b]}$ for the "b-th row" of P in section 4.3.)

Lemma 5.5. *Every Σ_1^B-Horn formula is provably equivalent in V_1-Horn to a Σ_1^B-Horn formula with a single existential quantifier. Specifically, if ϕ is Π_1^b-Horn with respect to $P_1, ..., P_k$ then V_1-Horn proves*

$$\exists P_1 ... \exists P_m \phi(P_1, ..., P_m) \leftrightarrow \exists P \phi(P^{[1]}, ..., P^{[m]})$$

Proof. For the left-to-right direction, use binary comprehension (Lemma 3.5) to define P satisfying

$$P(i, x) \leftrightarrow (i = 1 \wedge P_1(x)) \vee ... \vee (i = m \wedge P_m(x))$$

For the other direction, for $i = 1, ..., m$ use Σ_1^B-Horn comprehension to define P_i such that $P_i(x) \leftrightarrow P(i, x)$. □

Thus it suffices to prove the Theorem for Σ_1^B-Horn formulas of the form

$$\Phi \equiv \exists P \forall x_1 \leq t_1 ... \forall x_k \leq t_k \phi(\bar{x}, P) \tag{10}$$

where ϕ is Horn with respect to P.

The algorithm we wish to represent has two main steps (see the proof of Theorem 2.4): First create a propositional Horn formula $\textsc{Prop}^\Phi$ (which depends on the values for the free variables in Φ), and second apply the Horn Sat algorithm to determine whether $\textsc{Prop}^\Phi$ is satisfiable. We represent $\textsc{Prop}^\Phi$ using the arrays C, D, V, and we will present a Σ_1^B-Horn formula $\textsc{Prop}_\Phi(C, \tilde{C}, D, \tilde{D}, V, \tilde{V})$ which defines these arrays and their negations. Besides the indicated free variables, $\textsc{Prop}_\Phi$ also has as free variables the free variables of Φ. For the second step we present a Σ_1^B-Horn formula $\textsc{HornSat}(a, b, C, \tilde{C}, D, \tilde{D}, V, \tilde{V}, R, \tilde{R})$ (with all free variables indicated) which is independent of Φ and which sets the result variable $R(0)$ true iff $\textsc{Prop}^\Phi$ is satisfiable.

The arrays C, D, V together with the scalars a, b completely specify the formula $\textsc{Prop}^\Phi$ as follows. The atoms of $\textsc{Prop}^\Phi$ are $P(0), ..., P(a-1)$, and the clauses are $cl_0, ..., cl_{b-1}$. We allow both the empty clause and the special clause $\textsc{True}$. The arrays C, D, V are defined as follows: For $0 \leq x < b$, $0 \leq v < a$

- $C(x, v)$ asserts that clause cl_x contains the negative literal $\neg P(v)$.
- $D(x, v)$ asserts that clause cl_x contains the positive literal $P(v)$.
- $V(x)$ asserts that clause cl_x is the clause $\textsc{True}$.

Since $\textsc{Prop}^\Phi$ is a Horn formula, for each x, $D(x, v)$ can be true for at most one v.

The array bounds a, b are represented by terms $\hat{a}, \hat{b}$ in the free variables of Φ and are determined as follows. For each term s in $\phi(\bar{x}, P)$ in (10) let $\hat{s}$ be the result of replacing each variable $x_1, ..., x_k$ by its respective upper bound $t_1, ..., t_k$. Then the upper bound $\hat{a}$ on the arguments of $P()$ is

$$\hat{a} \equiv \hat{s}_1 + ... + \hat{s}_\ell$$

where $s_1, ..., s_\ell$ is a list of all terms such that $P(s_i)$ or $\neg P(s_i)$ occurs in Φ.

The upper bound $\hat{b}$ on the number of clauses in $\textsc{Prop}^\Phi$ is

$$\hat{b} \equiv \langle t_1, ..., t_s, m\rangle$$

where $t_1, ..., t_s$ are as in (10), m in the number of clauses in $\phi(\bar{x}, P)$, and $\langle ...\rangle$ is the tupling function (2).

Using the abbreviation

$$\bar{Q} \equiv C, \tilde{C}, D, \tilde{D}, V, \tilde{V}$$

we can now choose $\textsc{Run}_\Phi(R, \tilde{R})$ to be a Σ_1^B-Horn formula such that

$$\textsc{Run}_\Phi(R, \tilde{R}) \;\leftrightarrow\; \exists\bar{Q}[\textsc{Prop}_\Phi(\bar{Q}) \wedge \textsc{HornSat}(\hat{a}, \hat{b}, \bar{Q}, R, \tilde{R})] \tag{11}$$

In fact we take $\textsc{Run}_\Phi(R, \tilde{R})$ to be a suitable prenex form of the right hand side.

5.1 Definition of $\mathbf{Prop}_\Phi(C, \tilde{C}, D, \tilde{D}, V, \tilde{V})$

Below we define three Σ_0^B formulas $\psi_C(x, v), \psi_D(x, v), \psi_V(x)$ which characterize the three arrays C, D, V.

Lemma 5.6. $\textsc{Prop}_\Phi(\bar{Q})$ *can be defined in such a way that* $\exists\bar{Q}\textsc{Prop}_\Phi(\bar{Q})$ *is* Σ_1^B*-Horn and* V_1*-Horn proves*

$$(i)\;\; \exists\bar{Q}\textsc{Prop}_\Phi(\bar{Q})$$

$$\begin{aligned}(ii)\;\; &\textsc{Prop}_\Phi(\bar{Q}) \;\rightarrow\; \forall v < \hat{a}\forall x < \hat{b}\\ &[(C(x, v) \leftrightarrow \psi_C(x, v)) \wedge (D(x, v) \leftrightarrow \psi_D(x, v)) \wedge (V(x) \leftrightarrow \psi_V(x))\\ &\wedge (\tilde{C}(x, v) \leftrightarrow \neg\psi_C(x, v)) \wedge (\tilde{D}(x, v) \leftrightarrow \neg\psi_D(x, v)) \wedge (\tilde{V}(x) \leftrightarrow \neg\psi_V(x))]\end{aligned}$$

Proof. We apply Lemma 4.4 once each for ψ_C, ψ_D, ψ_V with S in the Lemma taken to be C, D, V, respectively, to obtain three Σ_1^B-Horn formulas $\psi_C^*, \psi_D^*, \psi_V^*$, and then let $\textsc{Prop}_\Phi(\bar{Q})$ be a prenex form of their conjunction. □

To define ψ_C, ψ_D, ψ_V let the Horn formula $\phi(\bar{x}, P)$ in (10) be the conjunction of the clauses $CL_0, ..., CL_{m-1}$. For $j = 0, ..., m-1$ let $\phi_j(\bar{x})$ be the quantifier-free formula which results by deleting all literals involving P from CL_j. Then we define

$$\begin{aligned}\psi_V(x) \equiv\; &\forall x_1 \leq t_1, ..., \forall x_k \leq t_k\\ &[(x = \langle x_1, ..., x_k, 0\rangle \rightarrow \phi_0(\bar{x})) \wedge ... \wedge (x = \langle x_1, ..., x_k, m-1\rangle \rightarrow \phi_{m-1}(\bar{x}))]\end{aligned}$$

Now let $\mathcal{S}$ be the set of indices j such that the clause CL_j has a positive literal of the form $P(u)$, and let for $j \in \mathcal{S}$ let that literal be $P(u_j(\bar{x}))$. Then we define

$$\psi_D(x,v) \equiv \neg\psi_V(x) \wedge \exists x_1 \le t_1, ..., \exists x_k \le t_k \bigvee_{j \in \mathcal{S}} [x = \langle x_1, ..., x_k, j\rangle \wedge v = u_j(\bar{x})]$$

For $j = 0, ..., m-1$ let $\neg P(u_j^0), ..., \neg P(u_j^{n_j - 1)})$ be the literals involving $\neg P$ in CL_j. Then

$$\psi_C(x,v) \equiv \neg\psi_V(x) \wedge \exists x_1 \le t_1, ..., \exists x_k \le t_k \bigvee_{j=0}^{m-1} \bigvee_{i=0}^{n_j - 1} [x = \langle x_1, ..., x_k, j\rangle \wedge v = u_j^i(\bar{x})]$$

5.2 Definition of HornSat$(a, b, C, \tilde{C}, D, \tilde{D}, V, \tilde{V}, R, \tilde{R})$

Although the Horn satisfiability algorithm is easy to describe informally, it is not straightforward to formalize in V_1-Horn. The propositional Horn satisfiability problem is complete for **P**, [12], and hence cannot be represented by a Σ_0^B formula. We need a more general form of Lemma 4.4 which allows us to use a Σ_1^B-Horn formula to define an array representing a given Σ_0^B formula, now in the presence of complementary variables $U, \tilde{U}$ which we want to existentially quantify.

Lemma 5.7. *Let $\psi(\bar{y}, U)$ be a Σ_0^B formula which may have free variables not indicated, but does not involve any of the variables $S, \tilde{S}, \bar{W}, \tilde{U}$ and has no occurrence of $|U|$. Then there is a formula $\psi^*(\bar{b}, S, \tilde{S}, \bar{W}, U, \tilde{U})$ not involving $\bar{y}$ but which may have other variables of ψ not indicated and which is Π_1^b Horn with respect to $S, \tilde{S}, \bar{W}, U, \tilde{U}$ such that V_1-Horn proves the following:*

$$\begin{aligned} (i) &\quad \exists S \exists \tilde{S} \exists \bar{W} \psi^*(\bar{b}, S, \tilde{S}, \bar{W}, U, \tilde{U}) \\ (ii) &\quad \psi^*(\bar{b}, S, \tilde{S}, \bar{W}, U, \tilde{U}) \wedge \forall z < s(U(z) \leftrightarrow \neg\tilde{U}(z)) \\ &\qquad \rightarrow \ \forall \bar{y} < \bar{b}[(S(\bar{y}) \leftrightarrow \psi(\bar{y}, U)) \wedge (\tilde{S}(\bar{y}) \leftrightarrow \neg\psi(\bar{y}, U))] \end{aligned}$$

where the term s is a provable upper bound on all terms r such that $U(r)$ occurs in ψ. A similar statement applies more generally to formulas $\psi(\bar{y}, U_1, ..., U_\ell)$ where the arrays U_i may have various dimensions.

Proof. We proceed by induction on the number of quantifiers in ψ, as in the proof of Lemma 4.4. The induction step is the same as before, but the base case now becomes more interesting. In this case ψ is quantifier-free, and we observe that the formula $(S(\bar{y}) \leftrightarrow \psi(\bar{y}, U)$ can be put into a conjunctive normal form which is Horn with respect to $S, U, \tilde{U}$ by taking the original CNF and replacing each positive literal of the form $U(r)$ by $\neg\tilde{U}(r)$. A similar remark applies to the formula $(\tilde{S}(\bar{y}) \leftrightarrow \neg\psi(\bar{y}, U)$. □

The algorithm represented by HornSat$(a, b, C, \bar{Q}, R, \tilde{R})$ attempts to find a satisfying assignment to the Horn formula Prop$^\Phi$ described by the parameters a, b, C, D, V. This is done by filling in an array $T(t, v)$, where $T(t, v)$ is the truth value assigned to the atom $P(v)$ after step t, $0 \le t, v < a$. Initially $T(0, v)$ is false, and at step $t+1$ $T(t+1, v)$ sets each $P(v)$ true such that $P(v)$ occurs positively in some clause not satisfied after step t. Once $P(v)$ is set true, it is never changed to false.

The following Σ_0^B formulas describe the array T and its negation $\tilde{T}$. First, $\textsc{Init}$ initializes T.

$$\textsc{Init} \;\equiv\; \forall v < a(\tilde{T}(0,v) \wedge \neg T(0,v))$$

In general we need to define a Σ_0^B formula $\textsc{Step}(v, T^{[t]})$ which expresses the value of $T(t+1,v)$ in terms of the values $T^{[t]}$ of T at time t. We define $\textsc{Step}$ using the one-dimensional array T_1 for $T^{[t]}$. First we need to define $\textsc{ClauseSat}(x, T_1)$ which asserts that assignment T_1 satisfies clause cl_x in $\textsc{Prop}^{\Phi}$.

$$\textsc{ClauseSat}(x, T_1) \;\equiv\; V(x) \vee \exists v < a[(C(x,v) \wedge \neg T_1(v)) \vee (D(x,v) \wedge T_1(v))]$$

Now $\textsc{Step}(v, T_1)$ holds iff either $P(v)$ is true under T_1 or there is a clause not satisfied by T_1 which has a positive literal $P(v)$.

$$\textsc{Step}(v, T_1) \;\equiv\; T_1(v) \vee \exists x < b(\neg\textsc{ClauseSat}(x, T_1) \wedge D(x,v)) \tag{12}$$

Now we apply Lemma 5.7 taking ψ to be $\textsc{Step}$ and $\bar{U}$ to be C, D, V, T_1 to obtain the formula $\textsc{Step}^*(a, S, \tilde{S}, \bar{W}, \bar{Q}, T_1, \tilde{T}_1)$ which is Π_1^b-Horn with respect to all of its displayed second-order variables and for which V_1-Horn proves the following versions of (i) and (ii) in the Lemma.

$$(i)' \quad \exists S \exists \tilde{S} \exists \bar{W} \textsc{Step}^*(a, S, \tilde{S}, \bar{W}, \bar{Q}, T_1, \tilde{T}_1)$$

$$\begin{aligned} (ii)' \quad & \textsc{Step}^*(a, S, \tilde{S}, \bar{W}, \bar{Q}, T_1, \bar{T}_1) \wedge \textsc{Neg} \wedge \forall v < a(T_1(v) \leftrightarrow \neg\tilde{T}_1(v)) \\ & \to \; \forall v < a[(S(v) \leftrightarrow \textsc{Step}(v, T_1)) \wedge (\tilde{S}(v) \leftrightarrow \neg\textsc{Step}(v, T_1))] \end{aligned}$$

where we define $\textsc{Neg}$ by

$$\textsc{Neg}(a, b, \bar{Q}) \;\equiv\; \forall v < a \forall x < b[(C(x,v) \leftrightarrow \neg\tilde{C}(x,v)) \wedge (D(x,v) \leftrightarrow \neg\tilde{D}(x,v)) \wedge (V(x) \leftrightarrow \neg\tilde{V}(x))] \tag{13}$$

Next we use the following formula to define the array T, where we have substituted $T^{[t+1]}$ for S and $T^{[t]}$ for T_1 in $\textsc{Step}^*$.

$$\textsc{TDef}(a, b, \bar{Q}, T, \tilde{T}) \;\equiv\; \textsc{Init}(T, \tilde{T}) \wedge \forall t < a \exists \bar{W} \textsc{Step}^*(a, T^{[t+1]}, \tilde{T}^{[t+1]}, \bar{W}, T^{[t]}, \tilde{T}^{[t]}) \tag{14}$$

Lemma 5.8. *V_1-Horn proves*

$$(i) \quad \exists T \exists \tilde{T} \; \textsc{TDef}(a, b, \bar{Q}, T, \tilde{T})$$

$$\begin{aligned} (ii) \quad & \textsc{TDef}(a, b, \bar{Q}, T, \tilde{T}) \wedge \textsc{Neg} \\ & \to \; \forall t < a \forall v < a[(T(t+1,v) \leftrightarrow \textsc{Step}(v, T^{[t]})) \wedge (\tilde{T}(t+1,v) \leftrightarrow \neg\textsc{Step}(v, T^{[t]}))] \end{aligned}$$

Proof. To prove (i), let $\textsc{TDef}'$ be obtained from $\textsc{TDef}$ by replacing the bounded quantifier $\forall t < a$ in the above definition of $\textsc{TDef}$ by $\forall t < y$. Define

$$\Phi(y) \equiv \exists T \exists \tilde{T} \; \textsc{TDef}'(y, a, b, \bar{Q}, T, \tilde{T})$$

By the Replacement Lemma $\Phi(y)$ is equivalent to a Σ_1^B-Horn formula, so we may use the induction scheme for $\Phi(y)$. This will establish (i), which is simply $\Phi(a)$.

For the base case $y = 0$ we need only satisfy $\textsc{Init}$, so we use the comprehension scheme to define T to be identically false and $\tilde{T}$ to be identically true.

Now assume the induction hypothesis and suppose that $T, \tilde{T}$ satisfy the existential quantifiers in $\Phi(y)$. Let $S, \tilde{S}$ satisfy the existential quantifiers in $(i)'$ when $T_1, \tilde{T}_1$ are replaced by $T^{[y]}, \tilde{T}^{[y]}$. Use comprehension to define the arrays $T', \tilde{T}'$ by

$$T'(t,v) \leftrightarrow \begin{cases} T(t,v) & \text{if } t \leq y \\ S(v) & \text{if } t > y \end{cases}$$

and

$$\tilde{T}'(t,v) \leftrightarrow \begin{cases} \tilde{T}(t,v) & \text{if } t \leq y \\ \tilde{S}(v) & \text{if } t > y \end{cases}$$

It follows from $\Phi(y)$ and $(i)'$ that $T', \tilde{T}'$ satisfy the existential quantifiers in $\Phi(y+1)$.

To prove (ii) we first claim that V_1-Horn proves

$$\text{TD}\textsc{ef} \wedge \textsc{Neg} \rightarrow \forall t \leq a \forall v < a (T(t,v) \leftrightarrow \neg \tilde{T}(t,v)) \tag{15}$$

V_1-Horn proves the RHS by induction on t, assuming $\text{TD}\textsc{ef} \wedge \textsc{Neg}$. For the base case $t = 0$ this follows from $\textsc{Init}(T, \tilde{T})$. The induction step $t \rightarrow t+1$ follows from $(ii)'$ above with $T^{[t+1]}, \tilde{T}^{[t+1]}$ substituted for $S, \tilde{S}$ and $T^{[t]}, \tilde{T}^{[t]}$ substituted for $T_1, \tilde{T}_1$.

Now (ii) follows from (15) and $(ii)'$ with this same substitution. $\square$

Now we define $\textsc{Sat}(T_1)$ to assert that the truth assignment T_1 satisfies $\textsc{Prop}^\Phi$.

$$\textsc{Sat}(T_1) \;\equiv\; \forall x < b \, \textsc{ClauseSat}(x, T_1)$$

The next lemma asserts that if the formula $\textsc{Prop}$ is satisfied at step t, then it remains satisfied for each subsequent step.

Lemma 5.9. *V_1-Horn proves*

$$\text{TD}\textsc{ef} \wedge \textsc{Neg} \;\rightarrow\; [t \leq y \leq a \wedge \textsc{Sat}(T^{[t]}) \rightarrow \textsc{Sat}(T^{[y]})]$$

Proof. This follows by applying induction on y to the RHS using Lemma 5.8 (ii). $\square$

Let $\textsc{Sat}^*(b, S, \tilde{S}, \bar{W}, \bar{Q}, T_1, \tilde{T}_1)$ be the result of applying Lemma 5.7 to $\textsc{Sat}(y, T_1)$, where we have introduced the new variable y as a placeholder. Now we define $\textsc{HornSat}$ to assert that there are arrays $T, \tilde{T}$ which satisfy $\text{TD}\textsc{ef}$ and such that $R(0)$ is true iff the truth assignment T at step a satisfies $\textsc{Prop}^\Phi$. Thus

$$\textsc{HornSat}(a, b, \bar{Q}, R, \tilde{R}) \;\equiv\; \exists T \exists \tilde{T} [\text{TD}\textsc{ef}(a, b, \bar{Q}, T, \tilde{T}) \wedge \exists \bar{W} \textsc{Sat}^*(1, R, \tilde{R}, \bar{W}, \bar{Q}, T^{[a]}, \tilde{T}^{[a]})] \tag{16}$$

It is clear from Lemma 5.7 that we may assume that the only atomic subformulas involving R or $\tilde{R}$ in $\textsc{HornSat}$ are $R(0)$ and $\tilde{R}(0)$ (by replacing $R(y)$ by $R(0)$ and $\tilde{R}(y)$ by $\tilde{R}(0)$), as required by the statement of Theorem 5.1.

Lemma 5.10. *V_1-Horn proves* $\exists R \exists \tilde{R}$ $\textsc{HornSat}(a, b, \bar{Q}, R, \tilde{R})$.

Proof. This is immediate from Lemma 5.8 and Lemma 5.7 (i) applied to $\textsc{Sat}$. $\square$

5.3 Proof of Theorem 5.1

Part (i) asserts that V_1-Horn proves $\exists R \exists \tilde{R}\ \textsc{Run}_\Phi(R, \tilde{R})$, where $\textsc{Run}_\Phi$ is defined in (11). This follows immediately from Lemma 5.6 (i) and Lemma 5.10.

The proof of (ii) requires formalizing the correctness proof of the Horn Sat algorithm. Correctness asserts that assuming $\bar{Q}$ is a proper code for a Horn formula $\textsc{Prop}$, then $\textsc{HornSat}$ implies $R(0)$ iff $\textsc{Prop}$ is satisfiable. To clarify the formal statement of correctness we write $\textsc{Sat}(T_1)$ as $\textsc{Sat}(a, b, \bar{Q}, T_1)$ with all of its free variables indicated.

Lemma 5.11 (Correctness of $\textsc{HornSat}$). *V_1-Horn proves*

$$\begin{aligned}&\textsc{HornSat}(a, b, \bar{Q}, R, \tilde{R}) \wedge \textsc{Neg}\\ &\rightarrow\ (R(0) \leftrightarrow \exists T_1 \textsc{Sat}(a, b, \bar{Q}, T_1)) \wedge (\tilde{R}(0) \leftrightarrow \neg \exists T_1 \textsc{Sat}(a, b, \bar{Q}, T_1))\end{aligned}$$

Proof. Reasoning in V_1-Horn, assume the hypotheses $\textsc{HornSat}$ and $\textsc{Neg}$, and let $T, \tilde{T}, \bar{W}$ satisfy the existential quantifiers in the definition (16) of $\textsc{HornSat}$. By Lemma 5.7 (ii) applied to $\textsc{Sat}(y, a, b, \bar{Q}, T_1)$ (where we have added the new variable y as a placeholder) with R for S and $T^{[a]}$ for T_1 we have

$$\begin{aligned}(ii)''\quad &\textsc{Sat}^*(1, a, b, R, \tilde{R}, \bar{W}, \bar{Q}, T^{[a]}, \tilde{T}^{[a]}) \wedge \textsc{Neg} \wedge \forall z < a (T^{[a]}(z) \leftrightarrow \neg \tilde{T}^{[a]}(z))\\ &\rightarrow\ (R(0) \leftrightarrow \textsc{Sat}(T^{[a]})) \wedge (\tilde{R}(0) \leftrightarrow \neg \textsc{Sat}(T^{[a]}))\end{aligned}$$

By (15), (16) and the hypotheses to the Correctness Lemma we conclude the hypotheses to $(ii)''$ and hence we conclude

$$(R(0) \leftrightarrow \textsc{Sat}(T^{[a]})) \wedge \tilde{R}(0) \leftrightarrow \neg \textsc{Sat}(T^{[a]})) \tag{17}$$

From this we conclude $R(0) \rightarrow \exists T_1 \textsc{Sat}(T_1)$ thus establishing one direction each in the two equivalences on the RHS of the Correctness Lemma (since $(ii)'' \rightarrow (R(0) \leftrightarrow \neg \tilde{R}(0))$.

Showing the other direction amounts to showing that under our hypotheses, $\exists T_1 \textsc{Sat}(T_1) \rightarrow \textsc{Sat}(T^{[a]})$. In other words, we must show that if $\textsc{Prop}$ is satisfiable, then it is satisfied by the final truth assignment given by the the Horn Sat algorithm. Formally it suffices to show that V_1-Horn proves

$$\textsc{TDef} \wedge \textsc{Neg} \wedge \textsc{Sat}(T_1)\ \rightarrow\ \textsc{Sat}(T^{[a]}) \tag{18}$$

First we show that $T^{[a]}$ is contained in every truth assignment satisfying $\textsc{Prop}$.

Lemma 5.12. *V_1-Horn proves*

$$\textsc{TDef} \wedge \textsc{Neg} \wedge \textsc{Sat}(T_1)\ \rightarrow\ \forall t < a \forall v < a (T(t, v) \rightarrow T_1(v))$$

Proof. The RHS is proved by induction on t. The base case $t = 0$ is vacuous because the condition $\textsc{Init}(T, \tilde{T})$ in the definition (14) of $\textsc{TDef}$ implies $T^{[0]}$ is identically false.

For the induction step we apply Lemma 5.8 (ii) and the definition (12) of $\textsc{Step}(v, T^{[t]})$. Thus the only way that $T(t + 1, v)$ can hold but not $T(t, v)$ is if some clause cl_x is not satisfied by $T^{[t]}$ and contains a positive literal $P(v)$. But by the induction hypothesis and our assumption that T_1 satisfies cl_x we have $\neg \textsc{ClauseSat}(x, t^{[t]}) \rightarrow T_1(v)$. $\square$

Now if $\textsc{Sat}(T_1)$ but $\neg\textsc{Sat}(T^{[a]})$ then there is a clause cl_x such that $\textsc{ClauseSat}(x, T_1)$ but $\neg\textsc{ClauseSat}(x, T^{[a]})$. Hence by the above Lemma cl_x contains a positive literal $P(v)$ such that $\neg T(a, v)$. Thus V_1-Horn proves

$$\textsc{TDef} \wedge \textsc{Neg} \wedge \textsc{Sat}(T_1) \wedge \neg\textsc{Sat}(T^{[a]}) \ \rightarrow\ \exists v < a \neg T(a, v) \tag{19}$$

There are only a atoms $P(0), ..., P(a-1)$ to be set, and as long as at least one clause is not satisfied every step sets at least one atom. It follows that after a steps $T^{[a]}$ must be identically true, contradicting (19).

To formalize the last part of the argument we introduce in the next subsection a counting formula $\textsc{NumOnes}(a, y, X)$, which asserts that the number of true values among $X(0), ..., X(a-1)$ is at least y. Using results in that subsection we now claim that V_1-Horn proves

$$\textsc{TDef} \wedge \textsc{Neg} \wedge \neg\textsc{Sat}(T^{[a]}) \wedge \textsc{Sat}(T_1) \ \rightarrow\ \textsc{NumOnes}(a, t, T^{[t]}) \tag{20}$$

This follows by applying induction on t to the RHS, using Lemma 5.14 (i) for the basis $t = 0$. For the induction step $t \rightarrow t+1$ we use Lemma 5.15 with $T^{[t]}$ for X, $T^{[t+1]}$ for Y, and t for y, and Lemma 5.8 (ii). The existence of v such that $\neg T(t, v) \wedge T(t+1, v)$ follows from our assumptions $\neg\textsc{Sat}(T^{[a]})$ (and hence $\neg\textsc{Sat}(T^{[t]})$ by Lemma 5.9) and $\textsc{Sat}(T_1)$ using Lemmas 5.8 (ii) and 5.12.

Finally (18) follows from (20) (with $t = a$) together with Lemma 5.14 (ii) and (19). This completes the proof of Lemma 5.11. □

We can now complete the proof of Theorem 5.1 (ii). By the definition (11) of $\textsc{Run}_\Phi$ and Lemma 5.11 if suffices to show that V_1-Horn proves the following two formulas.

$$\textsc{Prop}_\Phi(\bar{Q}) \rightarrow \textsc{Neg}(\hat{a}, \hat{b}, \bar{Q}) \tag{21}$$

$$\textsc{Prop}_\Phi(\bar{Q}) \ \rightarrow\ [\Phi \leftrightarrow \exists T_1(\textsc{Sat}(\hat{a}, \hat{b}, \bar{Q}, T_1))] \tag{22}$$

That (21) is provable follows from the definition (13) of $\textsc{Neg}$ and Lemma 5.6 (ii).

To show (22) is provable we refer to the definition (10) of Φ and show that V_1-Horn proves

$$\textsc{Prop}_\Phi(\bar{Q}) \ \rightarrow\ \forall x_1 \leq t_1 ... \forall x_k \leq t_k[\phi(\bar{x}, P) \leftrightarrow \textsc{Sat}(\hat{a}, \hat{b}, \bar{Q}, P)] \tag{23}$$

Recall (see the proof of Lemma 5.6) that $\phi(\bar{x}, P)$ is the conjunction of the clauses $CL_0, ..., CL_{m-1}$. By Lemma 5.6 (ii) and the definitions of Ψ_C, Ψ_D, Ψ_V, V_1-Horn proves for $j = 0, ..., m-1$

$$\textsc{Prop}_\Phi(\bar{Q}) \ \rightarrow\ \forall \bar{x} \leq \bar{t}[CL_j(\bar{x}, P) \leftrightarrow \textsc{ClauseSat}(\langle \bar{x}, j\rangle, P)]$$

This establishes the right-to-left direction of the equivalence in (23). To establish the other direction we also need the fact that V_1-Horn proves (assuming $\textsc{Prop}_\Phi(\bar{Q})$) that if x is not of the form $\langle x_1, ..., x_k, j\rangle$ then $\Psi_V(x)$ and hence $V(x)$ and hence $\textsc{ClauseSat}(x, P)$.

5.4 Counting in V_1-Horn

The results in this subsection are needed to complete the proof of Lemma 5.11 (Correctness of $\textsc{HornSat}$).

We define a Σ_1^B-Horn formula $\textsc{NumOnes}(a, y, X)$ which asserts that the number of true values among $X(0), ..., X(a-1)$ is at least y. First we define a formula $\textsc{Count}(a, M, \tilde{M}, X)$ which is

Π_1^b-Horn with respect to $M, \tilde{M}$ and which defines complementary arrays $M, \tilde{M}$ so that for $t, y \leq a$, $M(t, y)$ holds iff the number of true values among $X(0), ..., X(t-1)$ is at least y. We give recurrence equations in the style of the definition of $\textsc{Parity}(X)$ given after Theorem 2.4.

$$\begin{aligned}\textsc{Count}(a, M, \tilde{M}, X) \equiv\ & \forall t \leq a \forall y \leq a \\ & \quad M(t,0) \wedge \neg\tilde{M}(t,0) \wedge \neg M(0, y+1) \wedge \tilde{M}(0, y+1) \\ & \wedge(\neg M(t, y+1) \vee \neg\tilde{M}(t, y+1)) \\ & \wedge(M(t,y) \wedge X(t) \rightarrow M(t+1, y+1)) \\ & \wedge(M(t, y+1) \rightarrow M(t+1, y+1)) \\ & \wedge(\tilde{M}(t,y) \rightarrow \tilde{M}(t+1, y+1)) \\ & \wedge(\tilde{M}(t, y+1) \wedge \neg X(t) \rightarrow \tilde{M}(t+1, y+1))\end{aligned}$$

Lemma 5.13. *V_1-Horn proves*

$$\begin{aligned}&(i)\ \exists M \exists \tilde{M} \textsc{Count}(a, M, \tilde{M}, X) \\ &(ii) \textsc{Count}(a, M, \tilde{M}, X) \rightarrow [t \leq a \rightarrow \forall y \leq a(M(t,y) \leftrightarrow \neg\tilde{M}(t,y))]\end{aligned}$$

Proof. Since (i) is a Σ_1^B-Horn formula we may use induction on a. When $a = 0$ we use comprehension to explicitly define M such that $M(0,0)$, $M(1,0)$, $\neg M(0,1)$, and $(M(1,1) \leftrightarrow X(0))$, and similarly for $\tilde{M}$. For the induction step $a \rightarrow a+1$ we use comprehension to define the new values of $M, \tilde{M}$ using the recursion equations and the old values given by the induction hypothesis, in the style of the proof of Lemma 5.8 (i).

The proof of (ii) uses the induction scheme applied to $\Phi(t)$, where $\Phi(t)$ is the RHS. □

This result allows us to use $\neg M$ and $\tilde{M}$ interchangeably, and we shall do this freely in what follows.

Now we give the definition

$$\textsc{NumOnes}(a, y, X) \equiv \exists M \exists \tilde{M}[\textsc{Count}(a, M, \tilde{M}, X) \wedge M(a, y)]$$

Lemma 5.14. *V_1-Horn proves the following:*

$$\begin{aligned}&(i)\ \ \textsc{NumOnes}(a, 0, X) \\ &(ii)\ \textsc{NumOnes}(a, a, X) \rightarrow \forall v < a X(v)\end{aligned}$$

Proof. (i) follows immediately from the definitions of $\textsc{NumOnes}$ and $\textsc{Count}$.

To prove (ii) we first show that V_1-Horn proves

$$\textsc{Count}(a, M, \tilde{M}, X) \rightarrow \forall y < a(t < y \rightarrow \neg M(t,y)) \tag{24}$$

This follows by induction on t applied to the RHS, using the definition of $\textsc{Count}$.

Next we show that V_1-Horn proves

$$\textsc{Count}(a, M, \tilde{M}, X) \wedge \neg X(v) \rightarrow [v < t \leq a \rightarrow \neg M(t,t)] \tag{25}$$

This also follows by induction on t applied to the RHS, using (24).

Now (ii) follows from (25) by setting $t = a$. □

We introduce the abbreviation

$$X \subseteq_a Y \;\equiv\; \forall y < a(X(y) \to Y(y))$$

Lemma 5.15. *V_1-Horn proves*

$$X \subseteq_a Y \wedge v < a \wedge \neg X(v) \wedge Y(v) \wedge y < a \to [\textsc{NumOnes}(a, y, X) \to \textsc{NumOnes}(a, y+1, Y)]$$

Proof. First we claim that V_1-Horn proves each of the following formulas using induction on t; the second uses the first.

$$X \subseteq_a Y \wedge \textsc{Count}(a, M, \tilde{M}, X) \wedge \textsc{Count}(a, M', \tilde{M}', Y) \to \forall y < a(t \leq a \wedge M(t, y) \to M'(t, y))$$

$$\begin{aligned} X \subseteq_a Y \wedge & \neg X(v) \wedge Y(v) \wedge \textsc{Count}(a, M, \tilde{M}, X) \wedge \textsc{Count}(a, M', \tilde{M}', Y) \\ & \to \; \forall y < a(v < t \leq a \wedge M(t, y) \to M'(t, y+1)) \end{aligned}$$

Now the lemma follows from Lemma 5.13 and the formula immediately above with $t = a$. $\square$

6 Equivalence of V_1-Horn, P-def and QPV

The first-order theory QPV (called PV1 in [15]) has function symbols for all polynomial-time computable functions, and the axioms include defining equations for these functions (based on Cobham's Theorem) and induction on the length of numbers. The theory has been extensively studied [5, 2, 8, 15, 7] and shown to robustly capture the notion of "polynomial-time reasoning". Zambella's [24] theory P-def is a second-order version of QPV, and can shown to be equivalent to QPV by the method of RSUV isomorphism (see [15]). Here we show that V_1-Horn is equivalent in power to P-def. This implies that V_1-Horn is equivalent in power to QPV, but is most likely not as powerful as S^1_2 (see Section 1). We begin by showing how to add function symbols to V_1-Horn.

6.1 Adding function symbols to V_1-Horn

In section 2 we defined the class **P** in our second-order setting to consist of all relations of the form $R(x_1, ..., x_k, Y_1, ..., Y_m)$ recognizable in time bounded by a polynomial in $(x_1, ..., x_k, |Y_1|, ..., |Y_m|)$. In the same spirit we now define the class FP to consist of all functions $F(x_1, ..., x_k, Y_1, ..., Y_m)$ computable in time bounded by a polynomial in $(x_1, ..., x_k, |Y_1|, ..., |Y_m|)$. There are two kinds of functions in FP; *string functions*, denoted by upper-case letters F, take second-order objects as values, and *number* functions, denoted by lower-case letters f, take first-order objects as values. As before, number values are expressed in unary notation when defining computation time.

We say that a function has *arity* $\langle k, m\rangle$ if it takes k number arguments and m string arguments.

It is convenient to represent a string function $F(\bar{x}, \bar{Y})$ by its *bit graph* $B_F(\bar{x}, \bar{Y})$, defined by the condition

$$B_F(i, \bar{x}, \bar{Y}) \Leftrightarrow F(\bar{x}, \bar{Y})(i)$$

That is, $B_F(i, \bar{x}, \bar{Y})$ holds iff the i-th bit of $F(\bar{x}, \bar{Y})$ is 1. The following characterization of FP is straightforward.

Lemma 6.1. *(i) A string function $F(\bar{x}, \bar{Y})$ is in* FP *iff $|F(\bar{x}, \bar{Y})|$ is bounded by a polynomial in $(\bar{x}, |\bar{Y}|)$ and its bit graph B_F is in $\mathbf{P}$.*
(ii) A number function $f(\bar{x}, \bar{Y})$ is in FP *iff $f(\bar{x}, \bar{Y}) = |F(\bar{x}, \bar{Y})|$ for some string function F in* FP.

We now define a conservative extension V_1-Horn(FP) of V_1-Horn by introducing function symbols for polynomial time functions with defining equations based on the above Lemma.

Definition 6.2 (Specification of V_1-Horn(FP))**.** *The language $\mathcal{L}_A^2$(FP) is the language $\mathcal{L}_A^2$ of V_1-Horn extended by new function symbols. We define function symbols, terms, formulas, and Σ_1^B-Horn formulas for V_1-Horn(*FP*) by simultaneous recursion as follows. In general $\bar{x} = x_1, ..., x_k$ and $\bar{Y} = Y_1, ..., Y_m$.*

(i) To every first-order term $\ell(\bar{x}, \bar{Y})$ and Σ_1^B-Horn formula $\Phi(i, \bar{x}, \bar{Y})$ we associate an arity $\langle k, m\rangle$ string function symbol F with defining formulas (renaming ℓ as ℓ_F and Φ as Φ_F)

$$|F(\bar{x}, \bar{Y})| \leq \ell_F(\bar{x}, \bar{Y}) \tag{26}$$

$$\forall i < \ell(\bar{x}, \bar{Y})[F(\bar{x}, \bar{Y})(i) \leftrightarrow \Phi_F(i, \bar{x}, \bar{Y})] \tag{27}$$

To every arity $\langle k, m\rangle$ string function symbol F we associate an arity $\langle k, m\rangle$ number function symbol f with defining formula

$$f(\bar{x}, \bar{Y}) = |F(\bar{x}, \bar{Y})| \tag{28}$$

(ii) First-order variables and 0 and 1 are first-order terms and second-order variables are second-order terms.

(iii) If t_1, t_2 are first-order terms then $t_1 + t_2$ and $t_1 \cdot t_2$ are first-order terms. If T is a second-order term then $|T|$ is a first-order term.

(iv) If $t_1, ..., t_k$ are first-order terms and $T_1, ..., T_m$ are second-order terms, and f and F are arity $\langle k, m\rangle$ number and string function symbols, respectively, then $f(\bar{t}, \bar{T})$ is a first-order term and $F(\bar{t}, \bar{T})$ is a second-order term.

(v) If s, t are first-order terms and T is a second-order term then $s = t, s \leq t$ and $T(t)$ are atomic formulas. Formulas are built from atomic formulas as in V_1-Horn using $\wedge, \vee, \neg$ and the first and second-order quantifiers.

(vi) Σ_1^B-Horn formulas are defined as in Definition 2.2, with *term* and *formula* understood in the present context, and with the restriction that no term may include any quantified second-order variable P_i as a proper subpart. (This generalizes the restriction that $|P_i|$ may not appear. However formulas $P_i(t)$ may appear for any term t satisfying this restriction.)

The axioms of V_1-Horn(FP) are the same as for V_1-Horn except that the comprehension scheme is generalized to allow comprehension for all Σ_1^B-Horn formulas of V_1-Horn(FP), and the defining formulas introduced in (i) for all function symbols are included.

We refer to function symbols F and f introduced by (i) as *derived* function symbols, to distinguish them from the original function symbols $0, 1, +, \cdot, |\ |$ of V_1-Horn. In reasoning about V_1-Horn(FP) it is useful to define the *rank* of each function symbol by assigning rank 0 to the original function symbols and in general assigning 1 + the maximum of the ranks of function symbols in ℓ_F and Φ_F to each function symbol F introduced by (i) above, and 1 + the rank of F for each function symbol f introduced by (i) above.

We claim that (a) every function symbol introduced by (i) represents a polynomial-time function, and (b) each Σ_1^B-Horn formula Φ of V_1-Horn(FP) represents a relation in $\mathbf{P}$. Claims (a) and (b) are proved simultaneously by induction on the rank of the function symbol introduced in (a), and the maximum of the ranks of the function symbols occurring in Φ for (b). The base case follows from Theorem 2.4, and for the induction step (a) follows from (b) and Lemma 6.1. To prove (b), we observe that the proof of the if direction of Theorem 2.4 still goes through. In particular, given values for the free variables $\bar{z}, \bar{Y}$ and the quantified variables $\bar{x}$ in (4), every first-order term can be evaluated to a number and every second-order term can be evaluated to a string, because the restriction in the definition (vi) of Σ_1^B-Horn insures that no term involves quantified second-order variables P_i.

It is not hard to check that the results in the previous two sections apply to V_1-Horn($\mathtt{FP}$) as well as to V_1-Horn. This is true in particular to the main theorem on $\textsc{Run}_\Phi$.

Theorem 6.3. *Theorem 5.1 on* $\textsc{Run}_\Phi$*, and its corollaries, apply to* V_1*-Horn(*$\mathtt{FP}$*). Any derived function symbol occurring in* $\textsc{Run}_\Phi$*,* $\textsc{Neg}_\Phi$*, etc. also occurs in* Φ*.*

Proof. The formula $\textsc{Run}_\Phi(R, \tilde{R})$ is constructed from the two formulas $\textsc{Prop}_\Phi$ and $\textsc{HornSat}$. The formula $\textsc{HornSat}$ describes the propositional Horn satisfiability algorithm, is independent of Φ, and is the same in the present context. The formula $\textsc{Prop}_\Phi$ describes the propositional version of Φ. This does depend on Φ but it is constructed in the present context exactly as before. □

The following lemma is needed for the proof of the theorem below.

Lemma 6.4 (Term Bounding)**.** *(Here all variables are fully indicated.) For each first-order term* $t(\bar{x}, \bar{Y})$ *of* V_1*-Horn(*$\mathtt{FP}$*) there is a first-order bounding term* $\ell_t(\bar{x}, \bar{y})$ *of* V_1*-Horn such that*

$$V_1\text{-}Horn(\mathtt{FP}) \;\vdash\; t(\bar{x}, \bar{Y}) \leq \ell_t(\bar{x}, |\bar{Y}|)$$

For each second-order term $T(\bar{x}, \bar{Y})$ *there is a first-order bounding term* $\ell_T(\bar{x}, \bar{y})$ *of* V_1*-Horn such that*

$$V_1\text{-}Horn(\mathtt{FP}) \;\vdash\; |T(\bar{x}, \bar{Y})| \leq \ell_T(\bar{x}, |\bar{Y}|)$$

Proof. The two assertions are proved simultaneously by double induction, first on the highest rank of any function symbol occurring in t or T, and second on the maximum nesting depth of derived function symbols in t and T. □

Theorem 6.5. *For every* Σ_1^B*-Horn formula* $\Phi'(\bar{x}, \bar{Y})$ *of* V_1*-Horn(*$\mathtt{FP}$*) there is a* Σ_1^B*-Horn formula* Φ *of* V_1*-Horn such that*

$$V_1\text{-}Horn(\mathtt{FP}) \;\vdash\; \Phi'(\bar{x}, \bar{Y}) \leftrightarrow \Phi(\bar{x}, \bar{Y})$$

Corollary 6.6 (Conservativity)**.** *Every theorem of* V_1*-Horn(*$\mathtt{FP}$*) in the language of* V_1*-Horn is a theorem of* V_1*-Horn.*

Proof. It suffices to show that every model M of V_1-Horn has an expansion M' to the language $\mathcal{L}_A^2(\mathtt{FP})$ which is a model of V_1-Horn($\mathtt{FP}$). To define M' it suffices to specify functions on the universes of M interpreting each function symbol F and f introduced in Definition 6.2 (i), in such a way that the defining formulas are satisfied. First note that the value of each first-order function f is uniquely specified by (28) as a first-order element of M (assuming that F has been specified). Next note that for each tuple of values for the arguments of F, (26,27) uniquely specify the value of $F(\bar{x}, \bar{Y})$ as a set of first-order elements of M. Further by the theorem, the formula Φ_F specifying the bit graph of F is equivalent to a Σ_1^B-Horn formula of V_1-Horn, and therefore by Σ_1^B-Horn comprehension this set of elements is realized in M as a second-order object. Finally the comprehension axioms for all Σ_1^B-Horn formulas of V_1-Horn(FP) are satisfied by M', by the Theorem. $\square$

Proof of the Theorem. The proof that each such Φ' can be converted to an appropriate Φ is carried out by triple induction, first on the highest rank r of any function symbol occuring in Φ', second on the maximum nesting depth d of derived functions in any term in Φ' containing a function symbol of rank r, and third on the number of such maximal terms occurring in Φ'. The base case, $r = 0$, is trivial since we may take $\Phi = \Phi'$. Now suppose $r > 0$ and let

$$\Phi'(\bar{x}, \bar{Y}) \;\equiv\; \exists P_1 ... \exists P_a \forall z_1 < t_1 ... \forall z_b < t_b \phi'(\bar{z}, \bar{P}, \bar{x}, \bar{Y}) \tag{29}$$

where ϕ' is a quantifier-free Horn formula satisfying the conditions in Definition 2.2. We may suppose that none of the quantifier bounding terms t_i contains a function symbol not in V_1-Horn since by the Term Bounding Lemma 6.4 we can replace $\forall x_i < t_i$ by its $\forall x_i < \ell_{t_i}$ and add the clause $x_i < t_i$ as a conjunct to ϕ'.

We may replace each occurrence $f(...)$ of a first-order derived function symbol f by its definition $|F(...)|$ increasing the rank or nesting depth of derived function symbols. Therefore we may assume that no first-order derived function symbol occurs in Φ'.

Let r be the maximum rank of any function symbol occurring in Φ', let d be the maximum nesting depth of derived function symbols in terms of rank r, and let T be a second-order term in Φ' containing a function symbol of rank r and let T have derived nesting depth d. Then T has the form $F(\bar{s}, \bar{S})$ where F is a second-order function symbol, $\bar{s}$ are first-order terms and $\bar{S}$ are second-order terms. There are two cases, depending on how T occurs in Φ':

Case I: T occurs in a term $|F(\bar{s}, \bar{S})|$.

Case II: T occurs in an atomic formula $F(\bar{s}, \bar{S})(t)$.

For Case I, suppose that $F(\bar{y}, \bar{Z})$ is defined from $\ell_F(\bar{y}, \bar{Z})$ and $\Phi_F(\bar{y}, \bar{Z})$ in (i) of Definition 6.2. Then according to the axioms of V_1-Horn, $|F(\bar{y}, \bar{Z})|$ is 1 + the largest $j < \ell_F(\bar{y}, \bar{Z})$ such that $\Phi_F(j, \bar{y}, \bar{Z})$, or 0 if no such j exists. Therefore

$$V_1\text{-Horn}(\mathtt{FP}) \vdash [i = |F(\bar{y}, \bar{Z})| \;\leftrightarrow\; \Psi(i, \bar{y}, \bar{Z})] \tag{30}$$

where $\Psi(i, \bar{y}, \bar{Z})$ is the formula

$$\begin{aligned} &(i = 0 \wedge \forall j < \ell_F(\bar{y}, \bar{Z}) \neg \Phi_F(j, \bar{y}, \bar{Z})) \\ &\vee \exists i' < i[i = i' + 1 \wedge \Phi_F(i', \bar{y}, \bar{Z}) \wedge \forall j < \ell_F(\bar{y}, \bar{Z})(i \leq j \supset \neg \Phi_F(j, \bar{y}, \bar{Z}))] \end{aligned}$$

Notice that by definition of *rank*, any function symbol occurring in Φ_F or ℓ_F has smaller rank than that of F, and therefore rank less than r. Therefore by Corollary 5.3 and Theorem 6.3, Ψ is provably equivalent to a Σ_1^B-Horn formula all of whose derived function symbols have rank less than r, and hence by the induction hypothesis provably equivalent to a Σ_1^B-Horn formula of V_1-Horn. Thus we may assume that $\Psi(i, \bar{y}, \bar{Z})$ is a Σ_1^B-Horn formula with no derived function symbol.

Define $\Psi'(i, \bar{x}, \bar{z}, \bar{Y}) \equiv \Psi(i, \bar{s}, \bar{S})$ where we have indicated all possible free variables of Ψ'. Then by (30)

$$V_1\text{-Horn}(\mathtt{FP}) \;\vdash\; [i = |F(\bar{s}, \bar{S})| \leftrightarrow \Psi'(i, \bar{x}, \bar{z}, \bar{Y})] \tag{31}$$

The derived nesting depth of terms in $\bar{s}, \bar{S}$ is less than that of $F(\bar{s}, \bar{S})$, and hence by the induction hypothesis we may assume that $\Psi'(i, \bar{x}, \bar{z}, \bar{Y})$ is a Σ_1^B-Horn formula with no derived function symbol.

We now apply Corollary 5.4 to $\Psi'(i, \bar{z})$ (that is, we don't change Ψ', but now only indicate the variables $i, \bar{z}$) to obtain a Σ_1^B-Horn formula $\textsc{Run}_{\Psi'(i,\bar{z})}(b, \bar{c}, R, \tilde{R}, \bar{x}, \bar{Y})$ satisfying the Corollary. Here b is a bounding variable for i, $\bar{c}$ are bounding variables for $\bar{z}$, and we have indicated the free variables $\bar{x}, \bar{Y}$ which $\textsc{Run}_{\Psi'(i,\bar{z})}$ inherits from Ψ'.

Referring to (29), let ϕ'_i be ϕ' with each occurrence of $|F(\bar{s}, \bar{S})|$ replaced by the variable i. Then by Corollary 5.4 and (31), noting that $\textsc{Run}_{\Psi'(i,\bar{z})}$ does not contain any of $i, \bar{z}, \bar{P}$ free,

$$V_1\text{-Horn}(\texttt{FP}) \ \vdash \ [\Phi'(\bar{x}, \bar{Y}) \leftrightarrow \Phi''(\bar{x}, \bar{Y})]$$

where $\Phi''(\bar{x}, \bar{Y})$ is the formula

$$\exists R \exists \tilde{R} \exists \bar{P} \forall \bar{z} < \bar{t} \forall i < \ell_F(\bar{s}, \bar{S})[\textsc{Run}_{\Psi'(i,\bar{z})}(\ell_F(\bar{s}, \bar{S}), \bar{t}, R, \tilde{R}, \bar{x}, \bar{Y}) \ \wedge \ (\neg R(i, \bar{z}) \vee \phi'_i(\bar{z}, \bar{P}, \bar{x}, \bar{Y})]$$

Note that Φ'' can be converted to an equivalent Σ_1^B-Horn formula by first putting it into a suitable prenex form and then putting a copy of the literal $\neg R(i, \bar{z})$ inside every clause of ϕ'_i to make the disjunction into a Horn formula. The resulting Σ_1^B-Horn formula has one fewer occurrence of a term of derived depth d containing a function symbol of rank r (since T was removed from ϕ' in forming ϕ'_i and $\textsc{Run}_{\Psi'(i,\bar{z})}$ has no derived function symbol). Hence by the induction hypothesis, Φ'' is provably equivalent to a Σ_1^B-Horn formula with no derived function symbol.

The proof for Case II is similar, but easier. By reasoning as before, we can find a Σ_1^B-Horn formula $\Psi'(i, \bar{x}, \bar{z}, \bar{Y})$ with no derived function symbol such that (analogously to (31))

$$V_1\text{-Horn}(\texttt{FP}) \ \vdash \ [F(\bar{s}, \bar{S})(i) \leftrightarrow \Psi'(i, \bar{x}, \bar{z}, \bar{Y})] \tag{32}$$

Again we apply Corollary 5.4 to $\Psi'(i, \bar{z})$ to obtain a Σ_1^B-Horn formula $\textsc{Run}_{\Psi'(i,\bar{z})}(b, \bar{c}, R, \tilde{R}, \bar{x}, \bar{Y})$ satisfying the corollary. Again referring to (29), let ϕ'_R be ϕ' with each positive occurrence of $F(\bar{s}, \bar{S})(t)$ replaced by $\neg\tilde{R}(t, \bar{z})$ and each occurrence of $\neg F(\bar{s}, \bar{S})(t)$ replaced by $\neg R(t, \bar{z})$. (In this way ϕ'_R is Horn with respect to $R, \tilde{R}$ in Definition 2.2.) Then by Corollary 5.4 and (32),

$$V_1\text{-Horn}(\texttt{FP}) \ \vdash \ [\Phi'(\bar{x}, \bar{Y}) \leftrightarrow \Phi''(\bar{x}, \bar{Y})]$$

where now $\Phi''(\bar{x}, \bar{Y})$ is the formula

$$\exists R \exists \tilde{R} \exists \bar{P} \forall \bar{z} < \bar{t}[\textsc{Run}_{\Psi'(i,\bar{z})}(\ell_F(\bar{s}, \bar{S}), \bar{t}, R, \tilde{R}, \bar{x}, \bar{Y}) \ \wedge \ \phi'_R(\bar{z}, \bar{P}, \bar{x}, \bar{Y})]$$

Again Φ'' can be converted to an equivalent Σ_1^B-Horn formula by putting it into a suitable prenex form, and hence by the induction hypothesis Φ'' is provably equivalent to a Σ_1^B-Horn formula with no derived function symbol. □

6.2 Specification of P-def

We present a version of Zambella's [24] P-def which fits our notation and axioms. It is the same in spirit to Zambella's system. The system P-def is obtained from a *Base Theory BT* by introducing function symbols for all functions in FP, based on Cobham's recursion-theoretic characterization of the polynomial-time computable functions.

The Base Theory BT has the language $\mathcal{L}_A^2(=)$, which is $\mathcal{L}_A^2$with second-order equality. System BT has the same terms and formulas as V_1-Horn, except that atomic formulas include equations $X = Y$ between second-order variables. The axioms of BT consist of the axioms B1,...,B13,L1,L2 of V_1-Horn, the axiom E of extensionality (below) and the comprehension scheme for Σ_0^B formulas.

$$\text{E}: \qquad X = Y \leftrightarrow [\ |X| = |Y| \wedge \forall i < |X|(X(i) \leftrightarrow Y(i))] \tag{33}$$

As mentioned in Section 2, the Σ_0^B formulas represent precisely the $\mathtt{AC}^0$ relations. Analogously to $\mathtt{FP}$, we define $\mathtt{FAC}^0$ to be those polynomially-bounded string and number functions whose bit graphs are $\mathtt{AC}^0$ relations. (The functions in $\mathtt{FAC}^0$ are termed *rudimentary* in [24].) After [24], we define the $\mathtt{R}$-def to be BT augmented with function symbols for functions in $\mathtt{FAC}^0$ and their defining formulas.

More precisely, the language of $\mathtt{R}$-def is $\mathcal{L}_A^2(=)$ augmented with new function symbols, which are defined by simultaneous recursion along with terms, formulas and Σ_0^B formulas, as in Definition 6.2 with the following changes. In (i), Σ_1^B-Horn formula is replaced with Σ_B^0 formula. In (v), we now allow $S = T$ as an atomic formula, where S, T are second-order terms. In (vi) we replace the definition of Σ_1^B-Horn formula by that of Σ_0^B formula, which is a bounded formula in the language of $\mathtt{R}$-def with no second-order quantifier.

The axioms of $\mathtt{R}$-def are the axioms B1,...,B13,L1,L2, and E, together with comprehension over the Σ_0^B formulas of $\mathtt{R}$-def and the defining formulas for all derived function symbols.

By an easier version of the proofs of Theorem 6.5 and Corollary 6.6 we can show that $\mathtt{R}$-def is a conservative extension of the Base Theory BT.

We next name a string function symbol $\textsc{Chop}$ of $\mathtt{R}$-def of arity $\langle 1, 1\rangle$, where $\textsc{Chop}(x, Y)$ is intended to be the initial segment of Y of length at most x. The defining equations of $\textsc{Chop}$ are

$$|\textsc{Chop}(x, Y)| \leq x$$
$$\forall i < x[\textsc{Chop}(x, Y)(i) \leftrightarrow Y(i)]$$

We define $\mathtt{P}$-def to be the extension of $\mathtt{R}$-def obtained by introducing new function symbols and their defining formulas as follows:

To every first-order term $\ell_F(z, \bar{x}, \bar{Y})$ of $\mathtt{P}$-def and function symbols G_F, H_F of $\mathtt{P}$-def of arities $\langle k-1, m\rangle, \langle k, m+1\rangle$ we associate an arity $\langle k, m\rangle$ string function F with defining formulas

$$F(0, \bar{x}, \bar{Y}) = \textsc{Chop}(\ell_F(0, \bar{x}, \bar{Y}), G(\bar{x}, \bar{Y})) \tag{34}$$
$$F(z+1, \bar{x}, \bar{Y}) = \textsc{Chop}(\ell_F(z, \bar{x}, \bar{Y}), H(z, \bar{x}, \bar{Y}, F(z, \bar{x}, \bar{Y}))) \tag{35}$$

In addition we allow new function symbols to be introduced as in (26,27,28), where now Φ_F is any Σ_0^B formula in the language of $\mathtt{P}$-def.

The axioms for $\mathtt{P}$-def are the same as for $\mathtt{R}$-def, except we include the defining formulas for the new function symbols, and Σ_0^B formulas allow the new function symbols.

We remark that (26,27,28) allow the introduction of a function symbol for the composition of other function symbols. For example, we could take $\Phi_F(i, \bar{x}, \bar{Y})$ to be $G(H(\bar{x}, \bar{Y}))(i)$.

6.3 Relating V_1-Horn and $\mathtt{P}$-def

Theorem 6.7. $\mathtt{P}$-*def is a conservative extension of* V_1-*Horn.*

The next two lemmas prove the two directions. The proofs of the lemmas and of Theorem 6.5 actually show how to translate V_1-Horn(FP) and P-def back and forth in such a way that V_1-Horn is fixed.

Lemma 6.8. *Every theorem of V_1-Horn is a theorem of* P*-def.*

Proof. It suffices to show that every Σ_1^B-Horn-COMP axiom is a theorem of P-def. Since P-def allows the Σ_0^B-COMP axioms, this amounts to showing that P-def proves that each Σ_1^B-Horn formula is equivalent to some Σ_0^B formula in the language of P-def. This can be done by defining function symbols in P-def for witnessing the second-order quantifiers in the Σ_1^B-Horn formula (1) and proving them correct. This amounts to describing the Horn satisfiability algorithm in P-def, or more precisely formalizing the proof of Theorem 5.1 (describing $\textsc{Run}_\Phi$) in P-def. We will not carry out the details here, since as mentioned in the beginning of this section of the power of QPV (and hence P-def) has been well established. □

Lemma 6.9. *Every theorem of* P*-def in the language of V_1-Horn is a theorem of V_1-Horn.*

Proof. First note that using the extensionality axiom E (33), every equation $S = T$ between second-order terms is provably equivalent in P-def to a Σ_0^B formula (denoted $E(S = T)$) not involving second-order $=$. Therefore we may assume that formulas in P-def do not involve such second-order equations.

Now we claim that for every derived function symbol F of P-def there is a function symbol F' of V_1-Horn(FP) which represents the same function, such that V_1-Horn(FP) proves the translation of the defining formula for F. The translation is carried out by replacing each function symbol G in the defining formula by its V_1-Horn(FP) counterpart G', and by replacing each second-order equation $S = T$ by $E(S = T)$. ¿From this property a simple model-theoretic argument shows that for every formula Φ of P-def, if Φ is a theorem of P-def then its translation Φ' is a theorem of V_1-Horn(FP). The Lemma follows.

We define the translation of F to F' by induction on the rank of F. If F is introduced in P-def by (26,27) where Φ_F is a Σ_0^B formula, then we introduce F' in V_1-Horn(FP) by (26,27) where $\ell_{F'}$ is ℓ'_F (the translation of ℓ_F into vhorn(FP)) and $\Phi_{F'}$ is a Σ_1^B-Horn formula equivalent to Φ'_F, using Corollary 4.5 and Theorem 6.3. If f is introduced in P-def by (28) then f' is introduced in V_1-Horn(FP) using (28) with F' for F.

Now suppose that F is introduced in P-def by (34,35). The idea is to fix the arguments $(z, \bar{x}, \bar{Y})$ of F and present a formula defining an array $P(i, y)$ (and its negative counterpart $\tilde{P}(i, y)$) giving the i-th bit of $F(y, \bar{x}, \bar{Y})$, $0 \le i < \ell'_F(y, \bar{x}, \bar{Y})$, $0 \le y \le z$, where ℓ'_F is the translation of ℓ_F as a term of V_1-Horn(FP). The formula will recursively define all values of $P(i, y)$ and $\tilde{P}(i, y)$ successively for $y = 0, 1, ..., z$. To give the step from y to $y + 1$ we must translate the formula $H(z, \bar{x}, \bar{Y}, Z)(i)$ into one which is "Horn with respect to Z". In what follows we will suppress the variables $\bar{x}, \bar{Y}$.

Applying Theorem 6.5, let $\Psi(i, y, Z)$ be a Σ_1^B-Horn formula of V_1-Horn equivalent to the formula $H'(y, \textsc{Chop}(\ell'_F(y), Z))(i)$. Next apply Corollary 5.4 to obtain the formula $\textsc{Run}_{\Psi(i)}(b, R, \tilde{R}, y, Z)$. Now apply Lemma 6.11 below to $\textsc{Run}_{\Psi(i)}$, using the bound $\ell'_F(y)$ for ℓ to obtain an equivalent Σ_1^B-Horn formula not involving $|Z|$. Further modify this formula by replacing each positive subformula of the form $Z(t)$ by $(t < \ell'_F(y) \wedge \neg\tilde{Z}(t))$ (distribute $\vee$ over $\wedge$ to keep the quantifier-free part in CNF) and each occurrence of the form $\neg Z(t)$ by $(\neg Z(t) \vee \ell'_F(y) < t)$. The result is a formula $\overline{\textsc{Run}}_{\Psi(i)}(b, R, \tilde{R}, y, Z, \tilde{Z})$ which is Σ_1^B-Horn with respect to $Z, \tilde{Z}$ whose truth is unchanged if Z is replaced by $\textsc{Chop}(\ell'_F(y), Z)$. Further, defining the hypothesis $\textsc{Hypo}(Z, \tilde{Z})$ to be the formula

$$\textsc{Hypo} \equiv \forall j < \ell'_F(y)(Z(j) \leftrightarrow \neg\tilde{Z}(j))$$

it follows by Corollary 5.4 that V_1-Horn($\mathtt{FP}$) proves

$$\textsc{Hypo} \to \exists R \exists \tilde{R}\ \overline{\textsc{Run}}_{\Psi(i)}(b, R, \tilde{R}, y, Z, \tilde{Z}) \tag{36}$$

$$[\textsc{Hypo} \wedge \overline{\textsc{Run}}_{\Psi(i)}(b, R, \tilde{R}, y, Z, \tilde{Z})] \\ \to\ \forall i < b[(R(i) \leftrightarrow H'(y, \textsc{Chop}(\ell'_F(y), Z))(i)) \wedge (\tilde{R}(i) \leftrightarrow \neg R(i))] \tag{37}$$

Referring to (26,27), we take the defining term $\ell_{F'}(z)$ for $F'(z)$ in V_1-Horn($\mathtt{FP}$) to be $\ell'_F(z)$, and the bit graph formula $\Phi_{F'}(i, z)$ for $F'(z)$ to be a suitable prenex form of

$$\Phi_{F'}(i, z) \equiv \exists P \exists \tilde{P}(P(i, z) \wedge \hat{\Phi}(z, P, \tilde{P}))$$

where $\hat{\Phi}$ is

$$\hat{\Phi}(z, P, \tilde{P}) \ \equiv \forall j < \ell'_F(0)[(P(j, 0) \leftrightarrow G'()(j)) \wedge (\tilde{P}(j, 0) \leftrightarrow \neg G'()(j))] \wedge \\ \forall y < z\ \overline{\textsc{Run}}_{\Psi(i)}(\ell'_F(y+1), P(*, y+1), \tilde{P}(*, y+1), y, P(*, y), \tilde{P}(*, y))]$$

where for example the notation $P(*, y+1)$ indicates that each occurrence of the form $R(t)$ in $\overline{\textsc{Run}}_{\Psi(i)}$ is replaced by $P(t, y+1)$.

It remains to show that the translations of (34,35) follow in V_1-Horn($\mathtt{FP}$) from (26,27). First note that $\textsc{Chop}' = \textsc{Chop}$, since the defining formulas for $\textsc{Chop}$ in $\mathtt{P}$-def are also in V_1-Horn($\mathtt{FP}$). Next note that by (26) for F', the RHS's of the translations of (34,35) can be replaced by the second argument of $\textsc{Chop}$ in each case; that is by $G'()$ and $H'(z, F'(z))$ respectively. Now (34) follows easily from the definition of $\hat{\Phi}(0, P\tilde{P})$.

To establish the translation of (35) we make a series of Claims.

Claim 1: V_1-Horn($\mathtt{FP}$) $\vdash\ \hat{\Phi}(z, P, \tilde{P}) \to \forall y \leq z\ \textsc{Hypo}(P(*, y), \tilde{P}(*, y))$

This follows using induction on z and (37).

Claim 2: (Uniqueness of P) V_1-Horn($\mathtt{FP}$) proves

$$[\hat{\Phi}(z, P, \tilde{P}) \wedge \hat{\Phi}(z, Q, \tilde{Q})] \to \forall y \leq z \forall i < \ell'_F(y)(P(i, y) \leftrightarrow Q(i, y))$$

Again this follows using induction on z and (37) and Claim 1.

Claim 3: V_1-Horn($\mathtt{FP}$) $\vdash\ \hat{\Phi}(z, P, \tilde{P}) \to \forall y \leq z \forall i < \ell'_F(y)[P(i, y) \leftrightarrow \Phi_{F'}(i, y)]$

The left-to-right direction of the equivalence is immediate from the definition of $\hat{\Phi}$. The right-to-left direction requires Claim 2.

Claim 4: V_1-Horn($\mathtt{FP}$) $\vdash \exists P \exists \tilde{P} \hat{\Phi}(z, P, \tilde{P})$

This follows using induction on z, (36), and Claim 1.

Claim 5: V_1-Horn($\mathtt{FP}$) $\vdash \forall i < \ell'_F(z)[\Phi_{F'}(i, z+1) \leftrightarrow H'(z, F'(z))(i)]$

The left-to-right direction follows from the definition of $\Phi_{F'}$, Claim 1, (37), and Claim 3. The right-to-left direction uses Claim 4 in addition.

Finally the translation of (35) follows immediately from Claim 5.

This completes the proof of Lemma 6.9, except for the Lemmas below. $\square$

Lemma 6.10. *If $\Phi(Z)$ is a Σ_1^B-Horn formula not involving $|Z|$, then* $\textsc{Run}_\Phi$ *does not involve $|Z|$.*

Proof. Inspection of the proof of Theorem 5.1 (in particular (11)) shows that all terms appearing in $\textsc{Run}_\Phi$ are constructed from variables and terms appearing in Φ, using $0, 1, +, \cdot$. $\square$

Lemma 6.11. *Let ℓ be a term not involving $|Z|$ and let $\Phi(Z)$ be a Σ_1^B-Horn formula. Then there is a Σ_1^B-Horn formula $\Psi(Z)$ not involving $|Z|$ such that V_1-Horn proves*

$$|Z| \leq \ell \supset [\Phi(Z) \leftrightarrow \Psi(Z)]$$

Proof. This argument is similar to Case II in the proof of Theorem 6.5. We can define the relation $i = |Z|$ by a Σ_0^B formula $B(i, Z)$ not involving $|Z|$ but using the upper bound ℓ on $|Z|$, so

$$V_1\text{-Horn} \vdash |Z| \leq \ell \supset [i = |Z| \leftrightarrow B(i, Z)] \tag{38}$$

Using Corollary 4.5 (or Corollary 5.3 and the Lemma above) we may assume that $B(i, Z)$ is Σ_1^B-Horn. Let $\Phi'(Z)$ be the formula

$$\exists R \exists \tilde{R} \forall i < \ell[\textsc{Run}_{B(i)}(R, \tilde{R}, Z) \wedge (\neg R(i) \vee \Phi_i(i, Z))]$$

where $\Phi_i(i, Z)$ is obtained from $\Phi(Z)$ by replacing each occurrence of $|Z|$ by i. Then by the above Lemma $\Phi'(Z)$ does not contain $|Z|$, and by Corollary 6.3 and (38) V_1-Horn proves

$$|Z| \leq \ell \supset [\Phi(Z) \leftrightarrow \Phi'(Z)]$$

It remains to show that $\Phi'(Z)$ is provably equivalent to a Σ_1^B-Horn formula $\Psi(Z)$ which does not introduce an occurrence of $|Z|$. We write $\Phi'(Z)$ as $\exists R \exists \tilde{R} \phi(R, \tilde{R}, Z)$ and apply Corollary 5.3 to $\phi(R, \tilde{R}, Z)$ to obtain a Σ_1^B-Horn formula ϕ' equivalent to ϕ in which no terms $|R|, |\tilde{R}|, |Z|$ are introduced and take $\Psi(Z)$ to be $\exists R \exists \tilde{R} \phi'(R, \tilde{R}, Z)$. We may assume Ψ is Σ_1^B-Horn by replacing any positive occurrence of R in ϕ' by $\neg \tilde{R}$ and any positive occurrence of $\tilde{R}$ by $\neg R$. $\square$

7 Finite Axiomatizability

Here we show that both V^0 and V_1-Horn are finitely axiomatizable, and that the $\forall\Sigma_1^B$ consequences of V_1-Horn and the $\forall\Sigma_1^b$ consequences of S_2^1 are each finitely axiomatizable.

Since V^0 defines the uniform $\mathtt{AC}^0$ functions, it seems plausible that V_1-Horn could be axiomatized by V^0 together with a formula expressing the comprehension axiom for some predicate which is complete for $\mathtt{P}$ under uniform $\mathtt{AC}^0$ reductions. Hence the finite axiomatizability of V_1-Horn should follow from that for V^0. In our proof of Theorem 7.5 below, that predicate is the Horn satisfiability problem, which is complete for $\mathtt{P}$ [12].

Theorem 7.1. *V^0 is finitely axiomatizable.*

Proof. We must show that all Σ_0^B-COMP axioms follow from finitely many theorems of V^0 (see section 3).

Let $2-BASIC^+$ (or simply B^+) denote the $2-BASIC$ axioms along with finitely many theorems of V^0 asserting basic properties of $+$ and $\cdot$ such as commutativity, associativity, distributive laws, and cancellation laws involving $+$, $\cdot$, and $\leq$. These can be proved from the $2-BASIC$ axioms by induction on Σ_0^B formulas, as discussed in Section 3.

It suffices to show that k-ary comprehension (8) for all Σ_0^B formulas follow from B^+ and finitely many such comprehension instances. We use the notation $\Phi[\bar{a},\bar{Q}](\bar{x})$ to indicate that the Σ_0^B formula Φ can contain the free variables $\bar{a},\bar{Q}$ in addition to $\bar{x}=x_1,...,x_k$. Then $COMP_\Phi(\bar{a},\bar{Q},\bar{b})$ denotes the comprehension formula

$$\exists Y \leq \langle b_1,...,b_k\rangle \forall x_1 < b_1 ... \forall x_k < b_k (Y(\bar{x}) \leftrightarrow \Phi(\bar{x})) \tag{39}$$

We will show that $COMP_\Phi$ for the following 12 formulas Φ will suffice.

$$\begin{array}{lll}
\Phi_1(x_1,x_2) & \equiv & \exists y \leq x_1(x_1 = \langle x_2, y\rangle) \\
\Phi_2(x_1,x_2) & \equiv & \exists z \leq x_1(x_1 = \langle z, x_2\rangle) \\
\Phi_3[Q_1,Q_2](x_1,x_2) & \equiv & \exists y \leq x_1(Q_1(x_1,y) \wedge Q_2(y,x_2)) \\
\Phi_4[a](x,y) & \equiv & y = a \\
\Phi_5[Q_1,Q_2](x,y) & \equiv & \exists z_1 \leq y \exists z_2 \leq y(Q_1(x,z_1) \wedge Q_2(x,z_2) \wedge y = z_1 + z_2) \\
\Phi_6[Q_1,Q_2](x,y) & \equiv & \exists z_1 \leq y \exists z_2 \leq y(Q_1(x,z_1) \wedge Q_2(x,z_2) \wedge y = z_1 \cdot z_2) \\
\Phi_7[Q_1,Q_2,c](x) & \equiv & \exists y \leq c(Q_1(x,y) \wedge Q_2(x,y)) \\
\Phi_8[Q_1,Q_2,c](x) & \equiv & \exists y_1 \leq c \exists y_2 \leq c(Q_1(x,y_1) \wedge Q_2(x,y_2) \wedge y_1 \leq y_2) \\
\Phi_9[X,Q,c](x) & \equiv & \exists y \leq c(Q(x,y) \wedge X(y)) \\
\Phi_{10}[Q](x) & \equiv & \neg Q(x) \\
\Phi_{11}[Q_1,Q_2](x) & \equiv & Q_1(x) \wedge Q_2(x) \\
\Phi_{12}[Q,c](x) & \equiv & \forall y \leq c Q(x,y)
\end{array}$$

In the following lemmas, we abbreviate $COMP_{\Phi_i}(...)$ by C_i.

Lemma 7.2. *For each $k \geq 2$ and $1 \leq i \leq k$ let*

$$\Psi_{ik}(y,z) \equiv \exists x_1 \leq y ... \exists x_{i-1} \leq y \exists x_{i+1} \leq y ... \exists x_k \leq y(y = \langle x_1,...,x_{i-1},z,x_{i+1},...,x_k\rangle)$$

Then

$$B^+, C_1, C_2, C_3 \vdash COMP_{\Psi_{ik}}$$

Proof. We proceed by induction on k. For $k=2$ we have $\Psi_{1,2} \equiv \Phi_1$ and $\Psi_{2,2} \equiv \Phi_2$. For $k>2$, recall $\langle x_1,...,x_k\rangle = \langle\langle x_1,...,x_{k-1}\rangle, x_k\rangle$. Thus $\Psi_{kk} \equiv \Phi_2$. For $1 \leq i < k$ use $COMP_{\Phi_3}$ with Q_1 defined by $COMP_{\Phi_1}$ and Q_2 defined by $COMP_{\Psi_{i,k-1}}$. □

Lemma 7.3. *Let $t(\bar{x})$ be a term which in addition to variables $\bar{x}$ may involve other variables $\bar{a},\bar{Q}$. Let $\Psi_t[\bar{a},\bar{Q}](\bar{x},y) \equiv y = t(\bar{x})$. Then*

$$B^+, C_1, ..., C_6 \vdash COMP_{\Psi_t}(\bar{a},\bar{Q},\bar{b},d)$$

Proof. By using algebraic theorems in B^+ we may suppose that $t(\bar{x})$ is a sum of monomials in $x_1,...,x_k$, where the coefficients are terms involving $\bar{a},\bar{Q}$. The case $t \equiv u$, where u does not involve any x_i is obtained from $COMP_{\Phi_4}$ with $a \leftarrow u$. The cases $t \equiv x_i$ are obtained from Lemma 7.2. We then build monomials using $COMP_{\Phi_6}$ repeatedly, and build the general case by repeated use of $COMP_{\Phi_5}$. □

Lemma 7.4. *Let $t_1(\bar{x}), t_2(\bar{x})$ be terms with variables among $\bar{x}, \bar{a}, \bar{Q}$. Suppose*

$$\begin{array}{lcl} \Psi_1[\bar{a},\bar{Q}](\bar{x}) & \equiv & t_1(\bar{x}) = t_2(\bar{x}) \\ \Psi_2[\bar{a},\bar{Q}](\bar{x}) & \equiv & t_1(\bar{x}) \leq t_2(\bar{x}) \\ \Psi_3[\bar{a},\bar{Q},X](\bar{x}) & \equiv & X(t_1(\bar{x})) \end{array}$$

Then $B^+, C_1, ..., C_9 \vdash COMP_{\Psi_i}$, for $i = 1, 2, 3$.

Proof. $COMP_{\Psi_1}(\bar{a}, \bar{Q}, \bar{b})$ follows from $COMP_{\Phi_7}(P_1, P_2, c, b)$ with for $i = 1, 2$, P_i defined from $COMP_{\Psi_{t_i}}$ in Lemma 7.3 with $d \leftarrow t_1(\bar{b}) + t_2(\bar{b}) + 1$, so

$$\forall \bar{x} < \bar{b} \forall y < t_1(\bar{b}) + t_2(\bar{b}) + 1 (P_i(\bar{x}, y) \leftrightarrow y = t_i(\bar{x}))$$

In $COMP_{\Phi_7}$ we take $c \leftarrow t_1(\bar{b})$ and $b \leftarrow \langle b_1, ..., b_k \rangle$. We proceed similarly for $COMP_{\Psi_2}$, using $COMP_{\Phi_8}$.

For $COMP_{\Psi_3}(\bar{a}, \bar{Q}, X, \bar{b})$ we use $COMP_{\Phi_9}(X, P, c, b)$ with $c \leftarrow t_1(\bar{b})$ and $b \leftarrow \langle b_1, ..., b_k \rangle$ and P defined from Lemma 7.3 similarly to P_1 above. □

Now we can complete the proof of the theorem. Lemma 7.4 takes care of the case when Φ is an atomic formula. Then by repeated applications of $COMP_{\Phi_{10}}$ and $COMP_{\Phi_{11}}$ we handle the case in which Φ is quantifier-free.

Now suppose $\Phi(\bar{x}) \equiv \forall y \leq t(\bar{x}) \phi(\bar{x}, y)$. We assume as an induction hypothesis that we can define Q satisfying

$$\forall \bar{x} < \bar{b} \forall y < t(\bar{b}) + 1 [Q(\bar{x}, y) \leftrightarrow (y \leq t(\bar{x}) \rightarrow \phi(\bar{x}, y)]$$

Then $COMP_{\Phi}(\bar{b})$ follows from $COMP_{\Phi_{12}}(Q, c, b)$ with $c \leftarrow t(\bar{b})$ and $b \leftarrow \langle b_1, ..., b_k \rangle$. □

Theorem 7.5. *V_1-Horn is finitely axiomatizable.*

Proof. It suffices to show that Corollary 5.4 (i) and (ii) can be proved for any Σ_1^B-Horn formula $\Phi(y)$ using finitely many theorems of V_1-Horn as axioms. We first will show how to do this for Theorem 5.1 (i) and (ii), and then explain how to modify the proof to get the corollary.

First note that for each Σ_1^B-Horn formula Φ we can define a version of PROP_Φ such that (i) and (ii) in Lemma 5.6 are theorems of V^0. Thus we include the finite set of axioms for V^0 from Theorem 7.1 among the finite axioms for V_1-Horn. The proof of Theorem 5.1 depends on Lemma 5.6 (which we have established) and some properties of HORNSAT. Since HORNSAT is independent of Φ, we can take these properties as axioms.

To generalize the proof of Theorem 5.1 in order to prove Corollary 5.4, we incorporate the variable y in $\Phi(y)$ as an argument of each of the arrays $C, D, V, \tilde{C}, \tilde{D}, \tilde{V}$ to define the formula $\text{PROP}_\Phi(y)$ in a modified Lemma 5.6. Then y is not free in $\text{PROP}_\Phi(y)$ (although it could be free in PROP_Φ). The definition (16) of HORNSAT is modified so that the parameter y is incorporated as an argument of each of the arrays $R, \tilde{R}, T, \tilde{T}$. Then Corollary 5.4 follows in the same way as Theorem 5.1. □

Theorem 7.6. *V_1-Horn is axiomatized by its $\forall\Sigma_1^B$ consequences.*

Proof. It suffices to show that each Σ_1^B-Horn comprehension axiom is a consequence of $\forall\Sigma_1^B$ theorems of V_1-Horn. First we show that the second-order quantifiers in Σ_1^B-Horn formulas (1) can be bounded. That is, for each Σ_1^B-Horn formula Φ there is a Σ_1^B formula Φ^B such that $\forall\Sigma_1^B V_1\text{-Horn} \vdash (\Phi \leftrightarrow \Phi^B)$. To construct Φ^B replace each second-order quantifier $\exists P$ in Φ by a bounded quantifier $\exists P \leq t$, where t is a provable upper bound on all terms u such that $P(u)$ occurs in Φ. The equivalence of Φ and Φ^B requires only Ψ-COMP instances for formulas Ψ with no second-order quantifiers, and these instances are $\forall\Sigma_1^B$ formulas.

The comprehension axiom (6) for $\Phi(z)$ follows from Corollary 5.4 (i) and (ii). The Σ_1^B form of (i) we need is

$$\exists R \leq y \exists \tilde{R} \leq y \ \textsc{Run}'_{\Phi(z)}(y, R, \tilde{R})$$

where $\textsc{Run}'_{\Phi(z)}$ has suitable bounds on its second-order quantifiers. For (ii) we do not need the clause involving $\tilde{R}$. If we replace Φ by Φ^B then a suitable prenex form of the result is $\forall\Sigma_1^B$. □

Corollary 7.7. *The $\forall\Sigma_1^B$ consequences of V_1-Horn are finitely axiomatizable. The $\forall\Sigma_1^b$ consequences of S_2^1 are finitely axiomatizable.*

Proof. The first sentence follows by compactness from Theorems 7.6 and 7.5. Since V^1 is $\forall\Sigma_1^B$ conservative over P-def [24], it follows from Theorem 6.7 that the $\forall\Sigma_1^B$ consequences of V^1 and of V_1-Horn are the same, and hence are finitely axiomatizable. The second sentence of the Corollary is equivalent to asserting that the $\forall\Sigma_1^B$ consequences of V^1 are finitely axiomatizable, by the RSUV isomorphism. □

References

[1] D. M. Barrington, N. Immerman, and H. Straubing. On uniformity within NC^1. *Journal of Computer and System Sciences*, 41(3):274 – 306, 1990.

[2] S. Buss. *Bounded Arithmetic*. Bibliopolis, Naples, 1986.

[3] S. Buss. Axiomatizations and conservation results for fragments of bounded arithmetic. *Contemporary Mathematics*, 106:57–84, 1990.

[4] S. Buss. Relating the bounded arithmetic and polynomial time hierarchies. *Annals of Pure and Applied Logic*, 75:67–77, 1995.

[5] S. A. Cook. Feasibly constructive proofs and the propositional calculus. In *Proceedings of the Seventh Annual ACM Symposium on Theory of Computing*, pages 83 –97, 1975.

[6] S. A. Cook. CSC 2429S: Proof Complexity and Bounded Arithmetic. Course notes, URL: "http://www.cs.toronto.edu/∼sacook/csc2429h", Spring 1998.

[7] S. A. Cook. Relating the provable collapse of P to NC^1 and the power of logical theories. *DIMACS series in Discrete mathematics and theoretical computer science*, 39:73–91, 1998.

[8] S. A. Cook and Alasdair Urquhart. Functional interpretations of feasibly constructive arithmetic. *Annals of Pure and Applied Logic*, 63(2):103 – 200, 1993.

[9] R. Fagin. Generalized first-order spectra and polynomial-time recognizable sets. *Complexity of computation, SIAM-AMC proceedings*, 7:43–73, 1974.

[10] E. Grädel. The Expressive Power of Second Order Horn Logic. In *Proceedings of 8th Symposium on Theoretical Aspects of Computer Science STACS '91, Hamburg 1991*, volume 480 of *LNCS*, pages 466–477. Springer-Verlag, 1991.

[11] E. Grädel. Capturing Complexity Classes by Fragments of Second Order Logic. *Theoretical Computer Science*, 101:35–57, 1992.

[12] R. Greenlaw, H. J. Hoover, and W. L. Ruzzo. *Limits to Parallel Computation*. Oxford University Press, 1995.

[13] N. Immerman. Relational queries computable in polytime. *Information and Control*, 68:86 –104, 1986.

[14] N. Immerman. *Descriptive complexity.* Springer Verlag, New York, 1999.

[15] J. Krajiček. *Bounded Arithmetic, Propositional Logic, and Complexity Theory.* Cambridge University Press, New York, USA, 1995.

[16] J. Krajiček, P. Pudlák, and G. Takeuti. Bounded arithmetic and the polynomial time hierarchy. *Annals of Pure and Applied Logic*, 52:143–153, 1991.

[17] Daniel Leivant. Characterization of complexity classes in higher-order logic. In *Proceedings of the Second Annual Conference on Structure in Complexity Theory*, pages 203–217, 1987.

[18] Daniel Leivant. Descriptive characterizations of computational complexity. *Journal of Computer and System Sciences*, 39:51–83, 1989.

[19] A. Razborov. An equivalence between second-order bounded domain bounded arithmetic and first-order bounded arithmetic. In P. Clote and J. Krajiček, editors, *Arithmetic, proof theory and computational complexity*, pages 247–277. Clarendon Press, Oxford, 1993.

[20] U. Schöning and R. Pruim. *Gems of theoretical computer science.* Springer, Berlin, 1998.

[21] L. J. Stockmeyer. The polynomial-time hierarchy. *Theoretical Computer Science*, 3:1–22, 1977.

[22] G. Takeuti. RSUV isomorphism. In P. Clote and J. Krajiček, editors, *Arithmetic, proof theory and computational complexity*, pages 364–386. Clarendon Press, Oxford, 1993.

[23] M. Vardi. Complexity of relational query languages. *Information and Control*, 68:137 –146, 1986.

[24] D. Zambella. Notes on polynomially bounded arithmetic. *The Journal of Symbolic Logic*, 61(3):942–966, 1996.