

Integer Arithmetic

Syntax and Semantics

The integer arithmetic (IA) is the first order theory of integer numbers.

The alphabet of the integer arithmetic consists of:

- function symbols $+$, \cdot , S (S is the successor function $n \mapsto n + 1$)
- constant symbol 0

The semantics of IA is defined in the structure $\mathfrak{N} = \langle \mathbb{N}, +, \cdot, n \mapsto n + 1 \rangle$.

Examples

- The order relation is defined as $x \leq y : \exists z . x + y = z$
- The set of even numbers is defined by $even(x) : \exists y . x = y + y$
- The divisibility relation is defined as $x|y : \exists z . y = xz$
- The set of prime numbers is defined by
 $prime(x) : \forall yz . x = yz \rightarrow (y = 1 \vee z = 1)$
- The Conjecture of Goldbach:
 $\forall x . 2 \leq x \wedge even(x) \rightarrow \exists y, z . prime(y) \wedge prime(z) \wedge x = y + z$

Peano Arithmetic

An *axiomatic theory* is a set of formulae in which truth is derived from a (possibly infinite) set of *axioms*, e.g. Euclid's geometry is an axiomatic theory.

1. $0 \neq S(x)$

2. $S(x) = S(y) \rightarrow x = y$

3. $x + 0 = x$

4. $x + S(y) = S(x + y)$

5. $x \cdot 0 = 0$

6. $x \cdot S(y) = x \cdot y + x$

7. $\varphi(0) \wedge \forall x . [\varphi(x) \rightarrow \varphi(S(x))] \rightarrow \forall x . \varphi(x)$

Notice that the last point defines an infinite number of axioms.

Presburger Arithmetic

Definition

PA is the additive theory of natural numbers $\langle \mathbb{N}, + \rangle$

The following relations are Presburger definable:

$$\text{even}(x) \quad : \quad \exists y . x = y + y$$

$$x \leq y \quad : \quad \exists z . x + z = y$$

$$\text{zero}(x) \quad : \quad \forall y . x \leq y$$

$$\text{one}(x) \quad : \quad \exists z . \text{zero}(z) \wedge \neg x = z \wedge \forall y . y = z \vee x \leq y$$

$$x \equiv_m y \quad : \quad \exists z . x \leq y \wedge y - x = mz \vee x > y \wedge x - y = mz$$

Quantifier Elimination in PA

1. **Eliminate the negations** Replace $\neg(t_1 = t_2)$ by $t_1 < t_2 \vee t_2 < t_1$, $\neg(t_1 < t_2)$ by $t_1 = t_2 \vee t_2 < t_1$, and $\neg(t_1 \equiv_m t_2)$ by $\bigvee_{i=1}^{m-1} t_1 \equiv_m t_2 + i$.

Then rewrite the formula into DNF, i.e. a disjunction of $\exists x . \beta_1 \wedge \dots \wedge \beta_n$, where each β_i is one of the following forms:

$$nx = u - t$$

$$nx \equiv_m u - t$$

$$nx < u - t$$

$$u - t < nx$$

Quantifier Elimination in PA

2. **Uniformize the coefficients of x** Let p be the least common multiple of the coefficients of x . Multiply each atomic formula containing nx by $\frac{p}{n}$. In particular, $nx \equiv_m u - t$ becomes $px \equiv_{\frac{p}{n}m} \frac{p}{n}(u - t)$.

Quantifier Elimination in PA

Eliminate the coefficients of x Replace all over the formula px by x and add the new conjunct $x \equiv_p 0$

Special case If $x = u - t$ occurs in the formula, eliminate directly x by replacing it with $u - t$.

Quantifier Elimination in PA

Assume $x = u - t$ does not occur. We have a formula of the form

$$\exists x . \bigwedge_{j=1}^l r_j - s_j < x \wedge \bigwedge_{i=1}^k x < t_i - u_i \wedge \bigwedge_{i=1}^n x \equiv_{m_i} v_i - w_i$$

Let $M = [m_i]_{i=1}^n$. The formula is equivalent to:

$$\bigvee_{j=1}^l \bigvee_{q=1}^M \left[\bigwedge_{i=1}^l r_i - s_i < (r_j - s_j) + q \wedge \bigwedge_{i=1}^k (r_j - s_j) + q < t_i - u_i \wedge \bigwedge_{i=1}^n (r_j - s_j) + q \equiv_{m_i} v_i - w_i \right]$$

Decidability of PA

The result quantifier elimination in a Presburger formula is equivalent to a disjunction of conjunctions of atomic propositions of the following forms:

$$\sum_{i=1}^n a_i x_i + b \geq 0$$
$$\sum_{i=1}^n a_i x_i + b \equiv_n m$$

PA is decidable

One-dimensional Integer Sets

p -ary Expansions

Given $n \in \mathbb{N}$, its p -ary expansion is the word $w \in \{0, 1, \dots, p-1\}^*$ such that $n = w(0)p^k + w(1)p^{k-1} + \dots + w(k)p^0$, denoted also by $(n)_p$.

Note that the most significant digit is $w(0)$.

Conversely, to any word $w \in \{0, 1, \dots, p-1\}^*$ corresponds its value $[w]_p = w(0)p^k + w(1)p^{k-1} + \dots + w(k)p^0$.

Notice that $[w]_p = [0w]_p = [00w]_p = \dots$, i.e. the leading zeros don't change the value of a word.

p -automata

We consider one-dimensional sequences $s : \mathbb{N} \rightarrow \mathbb{N}$.

$$\begin{array}{cccccccccccccccc} & & 1 & 2 & & 4 & & & & 8 & & & & & & & & 16 & & \\ p_2 & : & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \dots \end{array}$$

Definition 1 *Let $p \geq 2$ be an integer. A p -automaton is a complete DFA $A = \langle S, q_0, T, \Lambda \rangle$ over the alphabet $\{0, 1, \dots, p - 1\}$, whose states are labeled with numbers from \mathbb{N} by a function $\Lambda : S \rightarrow \mathbb{N}$.*

A p -automaton defines a function $f : \{0, 1, \dots, p - 1\}^ \rightarrow \mathbb{N}$.*

Notice that the final states of a p -automaton may be designated by Λ .

p -automata

Definition 2 A sequence s is said to be p -recognizable iff there exists a p -automaton $A = \langle S, q_0, T, \Lambda \rangle$ such that, for all $n \in \mathbb{N}$:

- $q_0 \xrightarrow{(n)_p} q$, and
- $\Lambda(q) = s(n)$

We will always assume that any p -automaton has a loop $q_0 \xrightarrow{0} q_0$.

p_2 is 2-recognizable.

p -definability

Consider the theory $\langle \mathbb{N}, +, V_p \rangle$, where $p \in \mathbb{N}$, and $V_p : \mathbb{N} \rightarrow \mathbb{N}$ is:

- $V_p(0) = 1$,
- $V_p(x)$ is the greatest power of p dividing x .

$P_p(x)$ is true iff x is a power of p , i.e. $P_p(x) : V_p(x) = x$.

$x \in_p y$ iff x is a power of p and x occurs in the p -expansion of y with coefficient j :

$$x \in_{j,p} y : P_p(x) \wedge [\exists z \exists t . y = z + j \cdot x + t \wedge z < x \wedge (x < V_p(t) \vee t = 0)]$$

p -definability

A sequence $s : \mathbb{N} \rightarrow \mathbb{N}$ is p -definable if, for each $v \in \text{rng}(s)$ there exists a first-order formula φ_v of $\langle \mathbb{N}, +, V_p \rangle$ such that:

$$s^{-1}(v) = \{n \in \mathbb{N} \mid \models \varphi_v(n)\}$$

In other words:

$$s(n) = v \iff \varphi_v(n) . \forall n \in \mathbb{N}$$

The sequence p_2 is 2-definable:

$$p_2^{-1}(1) = \{n \in \mathbb{N} \mid \models V_2(n) = n\}$$

$$p_2^{-1}(0) = \{n \in \mathbb{N} \mid \models V_2(n) \neq n\}$$

Multi-dimensional Integer Sets

p -recognizability and p -definability

Let $(u, v) \in (\{0, 1, \dots, p-1\}^2)^*$ be a word, where $u, v \in \{0, 1, \dots, p-1\}^*$, $|u| = |v|$.

A p -automaton is defined now over $(\{0, 1, \dots, p-1\}^2)^*$.

The definitions of p -recognizability and p -definability are easily adapted to the m -dimensional case.

p -recognizability and p -definability

Consider $t : \mathbb{N}^2 \rightarrow \{0, 1\}$ defined as $t(n, m) = 0$ iff for some $k \geq 0$, we have $(n)_2(k) = (m)_2(k) = 1$, and $t(n, m) = 1$ otherwise.

$\uparrow m$								
	1	0	0	0	0	0	0	0
	1	1	0	0	0	0	0	0
	1	0	1	0	0	0	0	0
	1	1	1	1	0	0	0	0
	1	0	0	0	1	0	0	0
	1	1	0	0	1	1	0	0
	1	0	1	0	1	0	1	0
	1	1	1	1	1	1	1	\xrightarrow{n}

p -recognizability and p -definability

Consider $t : \mathbb{N}^2 \rightarrow \{0, 1\}$ defined as $t(n, m) = 0$ iff for some $k \geq 0$, we have $(n)_2(k) = (m)_2(k) = 1$, and $t(n, m) = 1$ otherwise.

				$\uparrow m$								
				1	0	0	0	0	0	0		
				1	1	0	0	0	0	0		
				1	0	1	0	0	0	0	0	
$(5)_2 =$	1	0	0	1	1	1	1	0	0	0	0	
$(4)_2 =$	1	1	0	1	0	0	0	1	0	0	0	
				1	0	0	0	1	1	0	0	
				1	1	0	0	1	0	1	0	
				1	0	1	0	1	0	1	0	
				1	1	1	1	1	1	1	1	\xrightarrow{n}

p -recognizability and p -definability

Consider $t : \mathbb{N}^2 \rightarrow \{0, 1\}$ defined as $t(n, m) = 0$ iff for some $k \geq 0$, we have $(n)_2(k) = (m)_2(k) = 1$, and $t(n, m) = 1$ otherwise.

				$\uparrow m$								
				1	0	0	0	0	0	0		
				1	1	0	0	0	0	0		
				1	0	1	0	0	0	0		
$(4)_2 =$	1	0	0	1	1	1	1	0	0	0	0	
$(3)_2 =$	0	1	1	1	0	0	0	1	0	0	0	
				1	1	0	0	1	1	0	0	
				1	0	1	0	1	0	1	0	
				1	1	1	1	1	1	1	1	\xrightarrow{n}

p -recognizability and p -definability

The sequence t is 2-recognizable.

The sequence t is 2-definable:

$$t^{-1}(0) \quad : \quad \exists z . z \in_2 x \wedge z \in_2 y$$

$$t^{-1}(1) \quad : \quad \forall z . \neg(z \in_2 x) \vee \neg(z \in_2 y)$$

p -recognizability and p -definability

Theorem 1 *Let $M \subseteq \mathbb{N}^m$, $m \geq 1$ and $p \geq 2$. Then M is p -recognizable if and only if M is p -definable.*

From Automata to Formulae

- $x \in_{j,p} y$ iff x is a power of p and the coefficient of x in $(y)_p$ is j :

$$x \in_{j,p} y : P_p(x) \wedge [\exists z \exists t . y = z + j \cdot x + t \wedge z < x \wedge (x < V_p(t) \vee t = 0)]$$

- $\lambda_p(x)$ denotes the greatest power of p occurring in $(x)_p$ and $\lambda_p(0) = 1$.

$$\lambda_p(x) = y : (x = 0 \wedge y = 1) \vee [P_p(y) \wedge y \leq x \wedge \forall z . (P_p(z) \wedge y < z) \rightarrow (x < z)]$$

From Automata to Formulae

Let $A = \langle S, q_0, T, \Lambda \rangle$ be a p -automaton, with $\Lambda : S \rightarrow \{0, 1\}$.

Suppose $S = \{q_0, q_1, \dots, q_{l-1}\}$ and replace w.l.o.g. q_k by $e_k = \langle 0, \dots, 1, \dots, 0 \rangle \in \{0, 1\}^l$.

$\langle n_1, \dots, n_m \rangle \in M$ iff $\langle (n_1)_p, \dots, (n_m)_p \rangle \in \mathcal{L}(A)$ iff exists $\langle y_1, \dots, y_l \rangle$:

- $\langle (y_1)_p(0), \dots, (y_l)_p(0) \rangle = \langle 1, 0, \dots, 0 \rangle$:

$$\varphi_1 : \bigwedge_{j=1}^l 1 \in_{q_0(j),p} y_j$$

From Automata to Formulae

- $\langle (y_1)_p(k), \dots, (y_l)_p(k) \rangle$ is a final state of A , with $p^k \geq \max_{1 \leq j \leq k} \lambda_p(x_j)$:

$$\varphi_2 : \bigvee_{\Lambda(q)=1} \bigwedge_{j=1}^l z \in_{q(j),p} y_j$$

- for all $0 \leq i < k$,

$$\langle (y_1)_p(i), \dots, (y_l)_p(i) \rangle \xrightarrow{\langle (x_1)_p(i), \dots, (x_m)_p(i) \rangle} \langle (y_1)_p(i+1), \dots, (y_l)_p(i+1) \rangle:$$

$$\varphi_3 : \forall t . P_p(t) \wedge t < z \wedge$$

$$\bigwedge_{T(q,(a_1,\dots,a_m))=q'} \left[\bigwedge_{j=1}^l t \in_{q(j),p} y_j \wedge \bigwedge_{j=1}^m t \in_{a_j,p} x_j \rightarrow \bigwedge_{j=1}^l p \cdot t \in_{q'(j),p} y_j \right]$$

From Automata to Formulae

$$\Phi_A : \exists y_1 \dots \exists y_l \exists z . P_p(z) \wedge z \geq \max_{1 \leq j \leq m} \lambda_p(x_j) \wedge \varphi_1(y_1, \dots, y_l) \wedge \\ \varphi_2(y_1, \dots, y_l, z) \wedge \varphi_3(x_1, \dots, x_m, y_1, \dots, y_l, z)$$

From Formulae to Automata

Build automata for the atomic formulae $x + y = z$ and $V_p(x) = y$, then compose them with union, intersection, negation and projection.

Corollary 1 *The theories $\langle \mathbb{N}, + \rangle$ and $\langle \mathbb{N}, +, V_p \rangle$ are decidable.*

The Cobham-Semenov Theorem

Base Dependence

Definition 3 *Two integers $p, q \in \mathbb{N}$ are said to be multiplicatively dependent if there exist $k, l \geq 1$ such that $p^k = q^l$.*

Equivalently, p and q are multiplicatively dependent iff there exists $r \geq 2$ and $k, l \geq 1$ such that $p = r^k$ and $q = r^l$.

Base Dependence

Lemma 1 *Let $p, q \geq 2$ be multiplicatively dependent integers. Let $m \geq 1$ and $s : \mathbb{N}^m \rightarrow \mathbb{N}$ be a sequence. Then s is p -recognizable iff it is q -recognizable.*

p^k -definable \Rightarrow p -definable Let $\phi(x, y) : P_{p^k}(y) \wedge y \leq V_p(x)$.

We have $V_{p^k}(x) = y \iff \phi(x, y) \wedge \forall z . \phi(x, z) \rightarrow z \leq y$.

We have to define P_{p^k} in $\langle \mathbb{N}, +, V_p \rangle$.

Base Dependence

$$P_{p^k}(x) : P_p(x) \wedge \exists y . x - 1 = (p^k - 1)y$$

Indeed, if $x = p^{ak}$ then $p^k - 1 | x - 1$.

Conversely, if assume x is a power of p but not of p^k , i.e. $x = p^{ak+b}$, for some $0 < b < k$.

Then $x - 1 = p^b(p^{ak} - 1) + (p^b - 1)$, and since $p^k - 1 | x - 1$, we have $p^k - 1 | p^b - 1$, contradiction.

Base Dependence

p -definable $\Rightarrow p^k$ -definable

$$V_{p^k}(x) = V_{p^k}(p^{k-1}x) \quad \rightarrow \quad V_p(x) = V_{p^k}(x)$$

$$V_{p^k}(x) = V_{p^k}(p^{k-2}x) \quad \rightarrow \quad V_p(x) = pV_{p^k}(x)$$

...

$$V_{p^k}(x) = V_{p^k}(px) \quad \rightarrow \quad V_p(x) = p^{k-2}V_{p^k}(x)$$

$$\text{else} \quad V_p(x) = p^{k-1}V_{p^k}(x)$$

Theorem 2 (Cobham-Semenov) *Let $m \geq 1$, and $p, q \geq 2$ be multiplicatively independent integers. Let $s : \mathbb{N}^m \rightarrow \mathbb{N}$ be a sequence. If s is p -recognizable and q -recognizable, then s is definable in $\langle \mathbb{N}, + \rangle$.*

Semilinear Sets

Definitions

$L(C, P) = \{x_0 + x_1 + \dots + x_m \mid x_0 \in C, x_1, \dots, x_m \in P\}$ for some $C, P \in \mathbb{N}^n$,

An element $x \in L(C, P)$ is of the form $x = x_0 + \sum_{i=1}^m \lambda_i x_i$, where $x_0 \in C$, $\lambda_i \in \mathbb{N}$ and $x_i \in P$, for all $1 \leq i \leq m$.

A set $M \in \mathbb{N}^n$ is said to be *linear* if $M = L(c, P)$ for $c \in \mathbb{N}^n$ and finite $P \subseteq \mathbb{N}^n$.

A set $M \in \mathbb{N}^n$ is said to be *semilinear* if $M = L(C, P)$ for finite $C, P \subseteq \mathbb{N}^n$.

A function $f : \mathbb{N}^n \rightarrow \mathbb{N}^m$ is said to be *linear* if for all $x, y \in \mathbb{N}^n$ we have $f(x + y) = f(x) + f(y)$.

Preliminaries

If $u = \langle u_1, \dots, u_n \rangle, v = \langle v_1, \dots, v_n \rangle \in \mathbb{N}^n$, we define $u \leq v$ iff $u(i) \leq v(i)$ for all $1 \leq i \leq n$.

Lemma 2 *Each set of pairwise incomparable elements of \mathbb{N}^n is finite. In consequence, each set $M \subseteq \mathbb{N}^n$ has a finite number of minimal elements.*

Lemma 3 *Let $M \subseteq \mathbb{N}^n$ be a semilinear set and $f : \mathbb{N}^n \rightarrow \mathbb{N}^m$ be a linear function. Then $f(M) \subseteq \mathbb{N}^m$ is a semilinear set.*

Let $w \in \mathbb{Z}^n$, $u_i, v_j \in \mathbb{N}^n$ and $a_i, b_j \in \mathbb{Z}$, $1 \leq i \leq p$, $1 \leq j \leq q$. Then there exists finite number of *minimal* tuples $\langle a_1, \dots, a_p, b_1, \dots, b_q \rangle$ such that:

$$w = \sum_{i=1}^p a_i u_i - \sum_{j=1}^q b_j v_j$$

Closure Properties

Theorem 3 *The class of semilinear subsets of \mathbb{N}^n , $n \geq 1$ is effectively closed under union, intersection and projection.*

Let

$$A = \{ \langle y_1, \dots, y_p, z_1, \dots, z_q \rangle \mid x_0 + \sum_{i=1}^p y_i x_i = x'_0 + \sum_{i=1}^q z_i x'_i \}$$

and

$$B = \{ \langle y_1, \dots, y_p, z_1, \dots, z_q \rangle \mid \sum_{i=1}^p y_i x_i = \sum_{i=1}^q z_i x'_i \}$$

Let $f : \mathbb{N}^{p+q} \rightarrow \mathbb{N}^n$ defined as $f(\langle y_1, \dots, y_p, z_1, \dots, z_q \rangle) = \sum_{i=1}^p y_i x_i$.

f is a linear function and $X \cap X' = x_0 + f(A)$. We prove that A is semilinear.

Let C and P be the sets of minimal elements of A and $B \setminus 0^{p+q}$, respectively. We prove that $A = L(C, P)$.

“ \subseteq ” $\mathbf{y} \cdot \mathbf{z} \in A \Rightarrow \exists \mathbf{y}' \cdot \mathbf{z}' \in C . \mathbf{y}' \cdot \mathbf{z}' \leq \mathbf{y} \cdot \mathbf{z}$. Let $\mathbf{y}'' \cdot \mathbf{z}'' = \mathbf{y} \cdot \mathbf{z} - \mathbf{y}' \cdot \mathbf{z}'$

$$\begin{aligned}
 \sum_{i=1}^p y''_i x_i &= \sum_{i=1}^p (y_i - y'_i) x_i \\
 &= \sum_{i=1}^p y_i x_i - \sum_{i=1}^p y'_i x_i \\
 &= (x'_0 - x_0) + \sum_{i=1}^q z_i x'_i - [(x'_0 - x_0) + \sum_{i=1}^q z'_i x'_i] \\
 &= \sum_{i=1}^q (z_i - z'_i) x'_i \\
 &= \sum_{i=1}^q z''_i x'_i
 \end{aligned}$$

Hence $\mathbf{y}'' \cdot \mathbf{z}'' \in B$. Prove that each element of B is a sum of elements of P .

Semilinear sets = Presburger-definable sets

Theorem 4 (Ginsburg-Spanier) *The class of semilinear subsets of \mathbb{N}^n coincides with the class of Presburger definable subsets of \mathbb{N}^n .*

“ \subseteq ” Let $M = L(C, P) \subseteq \mathbb{N}^k$ be a semilinear set, with $C = \{c_1, \dots, c_n\} \subset \mathbb{N}^k$ and $P = \{p_1, \dots, p_m\} \subset \mathbb{N}^k$.

The Presburger formula defining M is:

$$\phi(x_1, \dots, x_k) : \exists y_1 \dots \exists y_m \cdot \bigvee_{i=1}^n \bigwedge_{j=1}^k x_j = c_i + \sum_{j=1}^m y_j p_j$$

Semilinear sets = Presburger-definable sets

“ \supseteq ” Let $\phi(x_1, \dots, x_k)$ be a Presburger formula, i.e. a disjunction of conjunctions of atomic propositions of the following forms:

$$\sum_{i=1}^n a_i x_i + b \geq 0$$
$$\sum_{i=1}^n a_i x_i + b \equiv_n m$$

Each atomic proposition describes a semilinear set, hence their intersections and unions are again semilinear sets.

Semilinear sets are p -definable for any $p \geq 2$.