Presburger Arithmetic & Quantifier Elimination

Goal: Get a fine grasp on LIS by studying a logic that goes beyond L.I.S: Presburger Arithmetic.

Presburger Arithmetic (PA) is a first order theory of integers where only addition is allowed, and 0 and 1 are the only constants

Syntax

Let VAR be a set of variables.

A PA term t is defined by the grammar below:

$$t ::= c \mid VAR \mid t + t \mid c \cdot t$$ where $c \in \mathbb{Z}$ is a constant.

A PA atomic formula compares two terms via $\leq$:

$$\alpha ::= t < t \mid t \leq t \mid t > t \mid t \geq t \mid t = t$$

A PA formula is constructed from atomic formulas using $\vee, \wedge, \neg, \exists$, or $\forall$

$$f ::= \alpha \mid f \vee f \mid f \wedge f \mid \neg f \mid \exists x.f \mid \forall x.f$$

where α refers to an atomic formula, and x to a variable.

We use $f \Rightarrow g$ as a shorthand for $\neg f \vee g$.

Recall: The variable $x \in VAR$ is free in ϕ, if no $\exists$ or $\forall$ quantific binds it.

We often write $\phi(x)$ to denote a formula ϕ that has a free var. x.

We also use vectors of variables, i.e. $\phi(x)$ where x is a vector of two var. instead of $\phi(x,y)$.

We sometimes write $x \in \mathbb{N}$ for a variable x, even though formally x is just a var. name.

For a PA formula $\phi(x)$ with free variables $x \in \mathbb{N}^d$, and assignments $y \in \mathbb{N}^d$, we write $\phi(y)$ to denote the formula where we replace the variables with their assignments.

We say that an assignment $y \in \mathbb{N}^d$ satisfies such a $\phi(x)$, if $\phi(y)$ holds.

We say that a formula ϕ with no free variables in it holds, iff

- $\phi = \psi \vee \Delta$ and ψ or Δ hold,
- $\phi = \psi \wedge \Delta$ and ψ and Δ hold,
- $\phi = \neg\psi$ and ψ does not hold,
- $\phi = \exists x.\psi(x)$ and there is a $y \in \mathbb{N}$ such that $\psi(y)$ holds,
- $\phi = \forall x.\psi(x)$ and for all $y \in \mathbb{N}$, $\psi(y)$ holds.

We write $sol(\phi(x)) = \{y \in \mathbb{N}^d \mid \phi(y) \text{ holds}\}$.

Examples: $\forall x.\forall y.\exists z.\ x = y \vee x < z < y$ does not hold! For $x=0$ and $y=1$, there is no z.

($x < z < y$ ↖ shorthand for $x < z \wedge z < y$)

$\exists z.\exists t.\ 10 = z + 2t$ holds, $z = 8$ and $t = 1$

For $\phi(x,y) = \exists z.\, \exists t.\, x+2y=10 \wedge x = 2z \wedge y = 2t$; $sol(\phi(x,y)) = \{(2,4),(6,2)\}$

We can even express maximality / minimality

$$\phi(x,y) = \exists z.\; x+2y=100 \;\wedge\; y=2z+1 \;\wedge$$
$$(\forall v.\, \forall u.\; (\exists k.\, v+2u=100 \wedge u=2k+1) \Rightarrow (v \leq x))$$

In other words, for $\psi(x,y) = \exists z.\, x+2y=100 \wedge y=2z+1$, $\phi(x,y) = \psi(x,y) \wedge (\forall u.\forall v.\, \psi(u,v) \Rightarrow (v \leq x))$

Then $sol(\phi(x,y)) = \{(98,1)\}$ because ϕ is only satisfied by a maximal assignment to x.

PA subsumes LIS, since for $A \in \mathbb{Z}^{\ell \times d}$ and $b \in \mathbb{Z}^{\ell}$,

$$\phi(x) = \bigwedge_{i \leq \ell} \sum_{j \leq d} A_{ij} \cdot x[j] \geq b[i] \quad \text{has } sol(\phi(x)) = sol(A \cdot x \geq b).$$

What kind of spaces can solutions to PA formulas express?
Suprisingly: $sol(\phi(x))$ is semilinear

Theorem: Let $\phi(x)$ be a PA formula. Then $sol(\phi(x))$ is semilinear.

(Proof follows from the result later this week)

Many theorems over integers are so strong, that the problem of determing whether a formula holds is undecidable:

Theorem: Peano Arithmetic (Presburger + Multiplication) is undecidable, even without $\forall$-quantifiers.
(Not handled here)

But, PA is decidable.

There are many algorithms, but we focus on the famous quantifier elimination algorithm.

Quantifier Elimination: We say that a FO theory admits quantifier elimination (QE), if for all formulas $\phi(x)$, there is an equivalent $\phi'(x)$ (i.e. for all y, $\phi(y)$ iff $\phi'(y)$) without quantifiers (quantifier free)
This elimination is effective, if $\phi'(x)$ can be computed given $\phi(x)$.

To show that a theory admits QE, it suffices to show that all formulas of the form $\phi(x) = \exists z.\, \psi(x,z)$ where $\psi(x,z)$ is quantifier-free. The argument is by induction on the formula length, with the following inductive case:

Assume Prenex N.F. $\longrightarrow \phi = \forall z.\, \Delta(x,z)$ or $\phi = \exists z.\, \Delta(x,z)$.

(i.e. $\underbrace{\forall z_1 \exists z_2 \ldots \forall z_k}_{\text{quantifiers only at the beginning.}} \cdot \Delta(z_1,\ldots,z_k)$)

Ind. Hypothesis: $\Delta(x,z) \rightarrow$ quantifier free, equivalent $\Delta'(x,z)$

↓

Use Δ' + Make $\forall$ into $\exists$:

$\phi = \neg\exists z.\, \underbrace{\neg\Delta'(x,z)}_{\text{quantifier-free}}$ or $\phi = \exists z.\, \Delta'(x,z)$

↓

Eliminate the $\exists z$ quantifier.

Unfortunately, PA does not admit QE in its original form:

$$\phi(x) = \exists z.\ x = 10 \cdot z \longrightarrow sol(\phi(x)) = \{0, 10, 20, \dots\}$$

Any formula without quantifies and with 1 variable either (i) has finitely many solutions or

(ii) admits all numbers exceeding a bound $B \in \mathbb{N}$.

But by adding a module operator to get PA_{mod}, i.e. where we allow atomic formulas of the form

$$\alpha ::= t < t \mid t \leq t \mid t > t \mid t \geq t \mid t = t \mid t \equiv k \pmod{c}$$

where $c \in \mathbb{N}_{\geq 1}$, $k \in \mathbb{N}$ constants,

the resulting theory admits QE.

Theorem: PA_{mod} admits an effective QE procedure.

Proof: Tomorrow

Presburger QE

Goal: Quantifier Elimination for PA_{mod}

Theorem: For PA_{mod}, there is a quantifier elimination procedure, such that given a PA_{mod} formula $\phi(x)$, it computes the quantifier-free equivalent $\phi'(x)$ in time $TOWER(O(n))$ where n is the size of $\phi(x)$. Size of $\phi'(x)$ is also bounded by this value.

$$TOWER(k) = 2^{2^{2^{\cdot^{\cdot^{2^k}}}}\}k \text{ mal}}$$

This bound is not optimal, but it is enough for our purposes.

As we previously discussed, it suffices to eliminate one $\exists$-quantifier.

We assume Disjunctive Normal Form (DNF) to avoid distribution blow ups.

(i.e. $\phi = \bigvee_i \bigwedge_j \Delta_{i,j}$ where $\Delta_{i,j}$'s are atoms / their negations)

Lemma: There is a $2^{2^{2^{O(n)}}}$-time algorithm that given $\phi(\vec{x}) = \exists z.\, \psi(x,z)$ (multiple variables), where $\psi(x,z)$ quantifier-free and in DNF, it computes a quantifier-free equivalent $\phi'(x)$ in DNF.

First, we organize the quantifier-free portion $\psi(x,z)$. Pe DNF,

$$\psi(x,z) = \bigvee_{i \le k} \bigwedge_{j \le \ell_i} \Delta_{i,j}(x,z) \quad \text{for } \Delta_{i,j}(x,z) \text{ that are atomic formulas or their negations.}$$

First we express all non modulo atoms using $\le$: i.e. instead of $2x+y = z$, $2x+y \le z \wedge z \le 2x+y$;
and instead of $z < 2x+y$, $z \le 2x+y-1$.

We can also formulate the negations of non-modulo atomic formulas as $\le$-comparasions

$$\neg\left(\sum_{i \le k} c_i \cdot t_i \le \sum_{i \le \ell} d_i \cdot s_i\right) \equiv \left(\sum_{i \le k} c_i \cdot t_i > \sum_{i \le \ell} d_i \cdot s_i\right) \equiv \left(\sum_{i \le \ell} d_i \cdot s_i + 1 \le \sum_{i \le k} c_i \cdot t_i\right)$$

Formulas of the form $\neg\left(\sum_{i \le k} c_i \cdot y_i \equiv k \pmod{\ell}\right)$ can be broken down into

$$\bigvee_{0 < j < \ell} \sum_{i \le k} c_i \cdot y_i \equiv k + j.$$

Bringing this to DNF once more by distributing $\vee$ and $\wedge$,

$$\psi(x,z) \equiv \bigvee_{i \le k} \bigwedge_{j \le \ell_i} \tau_{i,j}(x,z) \quad \text{where } \tau_{i,j}(x,z) \text{ are atomic formulas that use } \le \text{ or mod.}$$

(formulas are equivalent)

We incur an exponential blow up during the distribution.

$$\text{Then } \phi(x) \equiv \exists z. \bigvee_{i \le k} \bigwedge_{j \le \ell_i} \tau_{i,j}(x,z) \equiv \bigvee_{i \le k} \exists z. \bigwedge_{j \le \ell_i} \tau_{i,j}(x,z).$$

Thus it suffices to show the following

Lemma: There is a $2^{2^{O(n)}}$-time algorithm that given $\phi(\bar{x}) = \exists z. \bigwedge_{i \le k} \tau_i(x,z)$, where $\tau_i(x,z)$ is of the form $\sum_{j \le d} c_j \cdot (x,z)[j] \le b$ or $\sum_{j \le d} c_j \cdot (x,z)[j] \equiv b \pmod m$ for all $i \le k$, it computes a quantifier-free equivalent $\phi'(x)$ in DNF.

($\bar{x}$: multiple variables; (x,z) is a vec. of vars so $(x,z)[j]$ is the j'th elem. on the list)

Idea: Show that there is a finite set of linear combination of vars in x such that if there is a solution for z, then one of these combinations is also a solution. Then, we can produce one copy formula for each combination, where we replace z with this combination, and bring them together with an $\bigvee(\vee)$.

First, the idea on an example

Consider $\exists z.\ x + 2z \le y+1 \wedge y \le x+2z \wedge (3z - x \equiv 0 \pmod 5)$

The solutions are $\{(0,10),(0,9),(1,3),(1,4),(1,13),(1,14),\dots\}$

We want to describe this relation by a quantifier-free formula.

Goal 1: Unify the constraints z has to fulfill.

First, leave z alone with the same positive coefficient

$$\equiv \exists z.\ 2z \le y - x + 1 \wedge y - x \le 2z \wedge (3z \equiv x \pmod 5)$$

$$\equiv \exists z.\ 6z \le 3y - 3x + 3 \wedge 3y - 3x \le 6z \wedge (6z \equiv 2x \pmod 5)$$

But this is the same as

$$\equiv \exists z'.\ z' \le \underbrace{3y - 3x + 3}_{\text{upper bound on } z'} \wedge \underbrace{3y - 3x \le z'}_{\text{lowerbound on } z'} \wedge (z' \equiv 2x \pmod 5) \wedge \underbrace{(z' \equiv 0 \pmod 6)}_{\text{since } z' = 6z \text{ must hold for } z \in \mathbb{N}}$$

Since z' must be at least $3y - 3x$, we have $z' \in [3y-3x+0, 3y-3x+1, \dots\} \cup \{0,1,2,\dots\}$.

But choosing too large values would brush up against the upper bound.

(These are new because $z' \ge 0$ is always a lowerbound.)

Only reason we might want to increase the value of z is to fulfill the modulo constraints.

We already get all relevant modulo values by considering the first $lcm(5,6) = 30$ values.

So,

$$\equiv \bigvee_{0 \le a \le 30} 3y - 3x + a \le 3y - 3x + 3 \wedge 3y - 3x \le 3y - 3x + a \wedge (3y - 3x + a \equiv 2x \pmod 5) \wedge (3y - 3x + a \equiv 0 \pmod 6) \wedge (3y - 3x + a \ge 0)$$

$$\vee \bigvee_{0 \le a \le 30} a \le 3y - 3x + 3 \wedge 3y - 3x \le a \wedge (a \equiv 2x \pmod 5) \wedge (a \equiv 0 \pmod 6) \wedge (a \ge 0)$$

We are done. There are no more quantifiers.

Now, the general proof.

Proof: Let $\phi(x) = \exists z. \bigwedge_{i \le k} \tau_i(x,z)$ where $\tau_i(x,z)$ are all $\le$- or $\equiv$-constraints.

Let $x \in \mathbb{N}^d$, i.e. let ϕ have d free variables.
First, we transform the constraints so that there is a $c \in \mathbb{N}$ such that if the constraint refers to z, one side of the constraint is $c \cdot z$.

This can be done similarly to the example: Move z terms such that their coefficient is positive, and move all the other terms to the other side.

Let $c_1, \dots, c_k$ be the resulting coefficients of z. Then multiply the constraint $\sum_{i \le d} a_i \cdot x[i] \le c_i \cdot z$ or $\sum_{i \le d} a_i \cdot x[i] \equiv c_i \cdot z \pmod m$ by $\frac{lcm(c_1,\dots,c_k)}{c_i}$

We get

$$\phi(x) \equiv \exists z. \bigwedge_{i \le s} \langle \ell_i, x \rangle + \alpha_i \le K \cdot z \wedge \bigwedge_{i \le t} K \cdot z \le \langle u_i, x \rangle + \beta_i \wedge \bigwedge_{i \le r} K \cdot z \equiv \langle q_i, x \rangle + \gamma_i \pmod{m_i} \wedge \Delta(x)$$

($\langle \ell_i, x \rangle = \sum_{j \le d} v_i[j] \cdot x[j]$; the constraints independent of z.)

for some $(\ell_i)_{i \le s}$, $(u_i)_{i \le t}$, and $(q_i)_{i \le r}$ in $\mathbb{Z}^d$, $(\alpha_i)_{i \le s}$, $(\beta_i)_{i \le t}$, $(\gamma_i)_{i \le r}$ in $\mathbb{Z}$, and $K = lcm(c_1, \dots, c_k)$.

Then, we can quantify for $K \cdot z$ instead. Clearly,

(1) $$\phi(x) \equiv \exists z'. \bigwedge_{i \le s} \langle \ell_i, x \rangle + \alpha_i \le z' \wedge \bigwedge_{i \le t} z' \le \langle u_i, x \rangle + \beta_i \wedge \bigwedge_{i \le r} z' \equiv \langle q_i, x \rangle + \gamma_i \pmod{m_i} \wedge z' \equiv 0 \pmod K \wedge \Delta(x)$$

We assume that $\ell_1 = 0$ and $\alpha_1 = 0$ to ensure $0 \le z'$ is explicitly a lowerbound.
Unlike the example, we have multiple lower- and upper-bounds for z'.
However, this is not a problem: For any assignment of x, one of $\langle \ell_j, x \rangle + \alpha_j$ will be the highest lowerbound.

We "guess" this bound via a disjunction, and apply the reasoning from the example, and claim

(2) $$\phi(x) \equiv \bigvee_{j \le s} \bigvee_{0 \le h \le M} \bigwedge_{i \le s} \langle \ell_i, x \rangle + \alpha_i \le \langle \ell_j, x \rangle + \alpha_j + h \wedge \bigwedge_{i \le t} \langle \ell_j, x \rangle + \alpha_j + h \le \langle u_i, x \rangle + \beta_i \wedge \bigwedge_{i \le r} \langle \ell_j, x \rangle + \alpha_j + h \equiv \langle q_i, x \rangle + \gamma_i \pmod{m_i} \wedge \Delta(x) \wedge \langle \ell_j, x \rangle + \alpha_j + h \equiv 0 \pmod K \wedge 0 \le \langle \ell_j, x \rangle + \alpha_j + h$$

(important for positivity of z'.)

for $M = lcm(m_1, \dots, m_r, K)$.

We claim that the right hand sides of (1) and (2) are equivalent, which will conclude the proof (quantifier is eliminated, formula is in DNF and $2^{O(|\phi|)}$ large).

for all x, (2) $\Rightarrow$ (1): Clear, because we only restrict which choices of z' are allowed.

(1) $\Rightarrow$ (2). Let $y \in \mathbb{N}^d$ be an assignment to $x \in \mathbb{N}^d$ such that the right hand side of (1) is satisfied.

Let $w \in \mathbb{N}$ be the assignment for z' that makes the inner formula true.

Let $\langle \ell_j, y\rangle + \alpha_j = \max_i \langle \ell_i, y\rangle + \alpha_i \overset{(*)}{\geq} \langle \ell_1, y\rangle + \alpha_1 = 0$

Now let $0 \leq h \leq M$ with $h \equiv w - (\langle \ell_j, y\rangle + \alpha_j) \pmod M$.

We show that assigning $\langle \ell_j, y\rangle + \alpha_j + h$ to z' also makes the right hand side true.

This shows that the rhs of (1) holds for y, since one of the clauses is obtained by replacing z' with $\langle \ell_j, x\rangle + \alpha_j$ and adding a $\geq$ check (which holds by $(*)$).

We simply check all constraints

Indep of z': The part independent of z', $\Delta(y)$ clearly holds since rhs of (1) is true.

Lowe Bd: Let $i \leq s$. $\langle \ell_i, y\rangle + \alpha_i \leq \langle \ell_j, y\rangle + \alpha_j + h$ holds by construction

Upper Bd: Let $i \leq t$, $w \leq \langle u_i, y\rangle + \beta_i$ already holds. By the j-th lowerbound, $w \geq (\langle \ell_j, y\rangle + \alpha_j)$. This implies $h \leq w - (\langle \ell_j, y\rangle + \alpha_j)$ and thus $\langle \ell_j, y\rangle + \alpha_j + h \leq w \leq \langle u_i, y\rangle + \beta_i$.

Modulo: Let $i \leq r$. We have

$$\langle \ell_j, y\rangle + \alpha_j + h \equiv w \pmod M$$

by definition of h. Then, since $M = \operatorname{lcm}(m_1, \ldots, m_r)$, $M = q \cdot m_i$ for $q \in \mathbb{N}$ and thus

$$w = \langle \ell_j, y\rangle + \alpha_j + h + n \cdot q \cdot m_i \text{ for } n \in \mathbb{Z} \text{ and}$$

$$w \equiv \langle \ell_j, y\rangle + \alpha_j + h \pmod{m_i}.$$

Since $w \equiv \langle q_i, y\rangle + \alpha_j \pmod{m_i}$ by w satisfying the formula, we get,

$$\langle \ell_j, y\rangle + \alpha_j + h \equiv w \equiv \langle q_i, y\rangle + \alpha_j \pmod{m_i}.$$