A lower bound for the complexity of Presburger arithmetic


$\langle \mathbb{R}, +, \leq \rangle$ is the first-order theory over real numbers with addition.

**Theorem 1**

For all $n \in \mathbb{N}$ there is a formula $\text{prod}_n(x, y, z)$ in $\langle \mathbb{R}, + \rangle$ such that for real numbers $A, B, C$ we have

$\text{prod}_n(A, B, C) \text{ is true } \iff A \in \mathbb{N} \land A < 2^{2^n} \land AB = C$

Furthermore, the size of $\text{prod}_n(x, y, z)$ is linear in $n$. 
Proof of theorem 1

By induction. Base \( n = 0 \) obvious.

Step:

- \( x \in \mathbb{N} \) and \( x < 2^{2^{n+1}} \) iff there exists \( x_1, x_2, x_3, x_4 \in \mathbb{N} \) all smaller than \( 2^{2^n} \) such that \( x = x_1x_2 + x_3 + x_4 \).

- We have \( z = xy = x_1(x_2y) + x_3y + x_4y \). Therefore, \( \text{prod}_{n+1}(x, y, z) \) is equivalent to
  \[
  \exists u_1, \ldots, u_5, x_1, \ldots, x_4. \text{prod}_n(x_1, x_2, u_1) \land \text{prod}_n(x_2, y, u_2) \land \text{prod}_n(x_1, u_2, u_3) \land \text{prod}_n(x_3, y, u_4) \land \text{prod}_n(x_4, y, u_5) \land x = u_1 + x_3 + x_4 \land z = u_3 + u_4 + u_5
  \]

- Problem: The size of the formula grows exponentially.
  Solution: Formulae like \( \phi = \varphi(x_1, y_1) \land \varphi(x_2, y_2) \land \varphi(x_3, y_3) \) can be equivalently written as
  \[
  \forall x. \forall y(((x = x_1 \land y = y_1) \lor (x = x_2 \land y = y_2) \lor (x = x_3 \land y = y_3)) \implies \varphi(x, y))
  \]
  Then, the size of \( \text{prod}_{n+1} \) is the size of \( \text{prod}_n \) + some constant.
Lower bound on the complexity of $\langle \mathbb{R}, + \rangle$

**Theorem 2**

For all $n \in \mathbb{N}$ there is a formula $\text{pow}_n(x, y, z)$ in $\langle \mathbb{R}, + \rangle$ such that for integers $a, b, c$ with $0 \leq a, b^a, c < 2^{2^n}$, $\text{pow}_n(a, b, c)$ is true iff $b^a = c$. Furthermore, the size of $\varphi_n(x, y, z)$ is linear in $n$.

The proof is similar to the proof of Theorem 1.
Proof of Theorem 2

By induction, we construct formulae \( e_k(x, y, z, u, v, w) \) such that for integers \( a, b, c \) with \( 0 \leq a < 2^{2^k} \), \( 0 \leq b^a, c < 2^{2^n} \) and real numbers \( A, B \) and \( C \),
\( e_k(a, b, c, A, B, C) \) is true in \( \langle \mathbb{R}, + \rangle \) iff \( A \in \mathbb{N}, A < 2^{2^n}, b^a = c \) and \( AB = C \).

We have \( e_k(0, 1, 1, A, B, C) \) iff \( \text{prod}_n(A, B, C) \)

- Induction base \( k = 0 \): We choose \( e_0 \) as
  \(((x = 0 \land z = 1) \lor (x = 1 \land z = y)) \land \text{prod}_n(u, v, w)\)

- Induction step:
  - \( x \in \mathbb{N} \) and \( x < 2^{2^{k+1}} \) iff there exists \( x_1, x_2, x_3, x_4 \in \mathbb{N} \) all smaller than \( 2^{2^k} \) such that \( x = x_1x_2 + x_3 + x_4 \).
  - Now, \( y^x = (y^{x_1})^{x_2}y^{x_3}y^{x_4} \). \( y^{x_1} \) is expressed by a \( z_1 \) such that \( e_k(x_1, y, z_1, 0, 0, 0) \) etc. For a product we use \( e_k(0, 1, 1, u, v, w) \).
  - Therefore, we can write \( e_{k+1}(x, y, z, u, v, w) \) using \( e_k \). With the same trick as in the proof of theorem 1 we need only one occurrence of \( e_k \).

- Finally, we have \( \text{pow}_n(x, y, z) \) iff \( e_n(x, y, z, 0, 0, 0) \).
Theorem 3

There exists a formula \( s_n(x, y) \) in \( \langle \mathbb{R}, + \rangle \) which is true iff \( x \) and \( y \) are integers, \( x < 2^{2^n} \) and \( y < 2^n \) and the \((y + 1)\)st digit \( x(y) \) of \( x \) is 1. Furthermore, the size of \( s_n(x, y, z) \) is linear in \( n \).

Proof: Exercise using the preceding theorems.

- With a natural number \(< 2^{2^n} \) one can code all binary sequences of length \( 2^n \).
- One can encode all computations of Turing machines up to length \( 2^n \). If the Turing machine makes at most \( 2^n \) steps, then it is necessary to consider configurations of size up to \( 2^n \). A computation is then a binary sequence of size \( 2^n2^n = 2^{2n} \).
- Therefore \( \langle \mathbb{R}, + \rangle \) is NEXPTIME-hard.
Theorem 4

For all $n$ there is a formula $\varphi_n(x, y, z)$ in $\langle \mathbb{N}, + \rangle$ such that for positive integer numbers $A, B, C$ we have

$$\varphi_n(A, B, C) \text{ is true } \iff A = BC \land A, B, C < \prod_{\substack{p \text{ prime} \\ p < f(n+2)}} p$$

where $f(n) = 2^{2^n}$. Furthermore, the size of $\varphi_n(A, B, C)$ is linear in $n$.

Basic idea of the proof: multiplication is addition of exponents in the decomposition of an integer in its prime factors (which is unique). It follows from the Prime Number Theorem that

$$\prod_{\substack{p \text{ prime} \\ p < f(n+2)}} p \geq 2^{f(n)^2} \geq 2^{2^{2n}}$$

- One can encode all computations of Turing machines with length $2^{2^n}$.
- Therefore $\langle \mathbb{Z}, + \rangle$ is 2-NEXPTIME-hard.