# Parametric Recognizability over the Integers

**Jakub Kaszycki**, Alessio Mansutti, Mikhail Starchak

SAMSA 2026
June 19th, 2026

# Linear integer arithmetic

Consider the structure $(\mathbb{Z}; +, \leqslant)$, with usual operations. We call it the **linear integer arithmetic** or LIA.

# Linear integer arithmetic

Consider the structure $(\mathbb{Z}; +, \leqslant)$, with usual operations. We call it the **linear integer arithmetic** or LIA.

**Fact ([Pre29])**

*Let* $|_k(n)$ *hold if and only if* $k \mid n$.

*Every first-order formula in the structure* $(\mathbb{Z}; +, \leqslant, |_2, |_3, \dots)$ *is equivalent to a quantifier-free formula.*

# Linear integer arithmetic

Consider the structure $(\mathbb{Z}; +, \leqslant)$, with usual operations. We call it the **linear integer arithmetic** or LIA.

## Fact ([Pre29])

*Let $|_k(n)$ hold if and only if $k \mid n$.*

*Every first-order formula in the structure $(\mathbb{Z}; +, \leqslant, |_2, |_3, \dots)$ is equivalent to a quantifier-free formula.*

This means that linear integer arithmetic is well-behaved, in particular:

## Corollary

*Linear integer arithmetic has a sound and complete proof system.*

# Linear integer arithmetic

Consider the structure $(\mathbb{Z}; +, \leqslant)$, with usual operations. We call it the **linear integer arithmetic** or LIA.

**Fact ([Pre29])**

*Let $|_k(n)$ hold if and only if $k \mid n$.*

*Every first-order formula in the structure $(\mathbb{Z}; +, \leqslant, |_2, |_3, \dots)$ is equivalent to a quantifier-free formula.*

This means that linear integer arithmetic is well-behaved, in particular:

**Corollary**

*Linear integer arithmetic has a sound and complete proof system.*

This is not obvious—for example, the above does not hold for $(\mathbb{Z}; +, \cdot)$ [Gö31].

# Integers and natural numbers

**Fact (Folklore)**

*The structures* $(\mathbb{N}; +)$ *and* $(\mathbb{Z}; +, \leqslant)$ *are bi-interpretable.*

# Integers and natural numbers

**Fact (Folklore)**

*The structures* $(\mathbb{N}; +)$ *and* $(\mathbb{Z}; +, \leqslant)$ *are bi-interpretable.*

We will work in $(\mathbb{N}; +)$ for simplicity. Nevertheless, all that we will do can be repeated for $\mathbb{Z}$, using a generalization of two's complement representation [Boi21].

# Integers and natural numbers

**Fact (Folklore)**

*The structures* $(\mathbb{N}; +)$ *and* $(\mathbb{Z}; +, \leqslant)$ *are bi-interpretable.*

We will work in $(\mathbb{N}; +)$ for simplicity. Nevertheless, all that we will do can be repeated for $\mathbb{Z}$, using a generalization of two's complement representation [Boi21].

Given a base $k \geqslant 2$, we can encode a tuple $(x_1, \ldots, x_m) \in \mathbb{N}^m$ in the $\mathrm{lsdf}_k$ encoding as a string over the alphabet $\{0, 1, \ldots, k-1\}^m$ of form

$$\begin{bmatrix} x_{1,0} \\ \vdots \\ x_{m,0} \end{bmatrix} \begin{bmatrix} x_{1,1} \\ \vdots \\ x_{m,1} \end{bmatrix} \cdots \begin{bmatrix} x_{1,n} \\ \vdots \\ x_{m,n} \end{bmatrix},$$

where $x_{i,j}$ is the $j$-th least significant digit of $x_i$ and $x_1, \ldots, x_m < k^n$. This encoding is not unique.

# LIA and automata

**Fact ([Bü60])**

*Every set of tuples of natural numbers definable using a LIA formula forms a regular language in the* $\mathsf{lsdf}_k$ *encoding in any base* $k$.

# LIA and automata

**Fact ([Bü60])**

*Every set of tuples of natural numbers definable using a LIA formula forms a regular language in the $\mathrm{lsdf}_k$ encoding in any base $k$.*

**Fact ([GS66])**

*LIA definable sets are exactly the semilinear sets, that is sets of form*

$$(b_1 + P_1^*) \cup \cdots \cup (b_n + P_n^*)$$

*for $b_i \in \mathbb{N}^m$, $P_i \subseteq_{fin} \mathbb{N}^m$.*

# LIA and automata

## Fact ([Bü60])

*Every set of tuples of natural numbers definable using a LIA formula forms a regular language in the $lsdf_k$ encoding in any base $k$.*

## Fact ([GS66])

*LIA definable sets are exactly the semilinear sets, that is sets of form*

$$(b_1 + P_1^*) \cup \cdots \cup (b_n + P_n^*)$$

*for $b_i \in \mathbb{N}^m$, $P_i \subseteq_{fin} \mathbb{N}^m$.*

## Corollary

*The regular language $0^*10^*$ encodes the set $\{k^n \mid n \in \mathbb{N}\}$, which is not LIA definable for any base $k$.*

# LIA and automata

**Fact ([Bü60])**

*Every set of tuples of natural numbers definable using a LIA formula forms a regular language in the $lsdf_k$ encoding in any base $k$.*

**Fact ([GS66])**

*LIA definable sets are exactly the semilinear sets, that is sets of form*

$$(b_1 + P_1^*) \cup \cdots \cup (b_n + P_n^*)$$

*for $b_i \in \mathbb{N}^m$, $P_i \subseteq_{fin} \mathbb{N}^m$.*

**Corollary**

*The regular language $0^*10^*$ encodes the set $\{k^n \mid n \in \mathbb{N}\}$, which is not LIA definable for any base $k$.*

Is there an arithmetic that captures all regular languages?

# Büchi-Bruyère arithmetic

Let $V_k(n) = k^m$ such that $k^m \mid n$ and $k^{m+1} \nmid n$.

**Fact ([Bü60, BHMV94])**

*The lsdf$_k$ encodings of sets of m-tuples definable in $(\mathbb{N}; +, V_k)$ are exactly the regular languages over an alphabet of size $k^m$ closed under appending trailing zeroes.*

Let us call this structure $\mathrm{BA}_k$.

# Büchi-Bruyère arithmetic

Let $V_k(n) = k^m$ such that $k^m \mid n$ and $k^{m+1} \nmid n$.

**Fact ([Bü60, BHMV94])**

*The lsdf$_k$ encodings of sets of m-tuples definable in $(\mathbb{N}; +, V_k)$ are exactly the regular languages over an alphabet of size $k^m$ closed under appending trailing zeroes.*

Let us call this structure BA$_k$. For every $k$, the theory is different. In particular:

**Fact ([Cob69])**

*If k and l are two numbers that do not share a common power, any set definable both in BA$_k$ and BA$_l$ is definable in LIA.*

# Theory of $BA_k$

Büchi-Bruyère arithmetic is also well-behaved, in particular:

**Corollary**

*$BA_k$ has a sound and complete proof system.*

# Theory of $BA_k$

Büchi-Bruyère arithmetic is also well-behaved, in particular:

**Corollary**

*$BA_k$ has a sound and complete proof system.*

**Fact ([Bü60, BHMV94])**

*The theory of $BA_k$ is TOWER-complete.*

# Theory of $\mathsf{BA}_k$

Büchi-Bruyère arithmetic is also well-behaved, in particular:

## Corollary

*$BA_k$ has a sound and complete proof system.*

## Fact ([Bü60, BHMV94])

*The theory of $BA_k$ is TOWER-complete.*

## Fact ([HR21])

*Every first-order formula of $BA_k$ is equivalent to a $\Sigma_2$ formula.*

*For any $k$, there exists a first-order formula with one free variable that is not equivalent to any $\Sigma_1$ formula.*

# A parametric view

Cobham's theorem looks at the formulae of $\mathsf{BA}_k$ defining a fixed set for different bases.

# A parametric view

Cobham's theorem looks at the formulae of $\mathrm{BA}_k$ defining a fixed set for different bases.

What if we look at the sets defined by a fixed formula over the common signature $\mathbb{N} \cup \{+, V\}$, interpreted in different bases? Let us call this PBA.

# A parametric view

Cobham's theorem looks at the formulae of $\mathrm{BA}_k$ defining a fixed set for different bases.

What if we look at the sets defined by a fixed formula over the common signature $\mathbb{N} \cup \{+, V\}$, interpreted in different bases? Let us call this PBA.

This might result in infinitely many different sets. How to encode them succinctly?

# Symbolic automata

## Definition (Symbolic automaton — intuition)

A **symbolic automaton** is an automaton with formulae on transitions.

# Symbolic automata

## Definition (Symbolic automaton — intuition)

A **symbolic automaton** is an automaton with formulae on transitions. Actually, there is much more to them [DV21].

# Symbolic automata

## Definition (Symbolic automaton — intuition)

A **symbolic automaton** is an automaton with formulae on transitions. Actually, there is much more to them [DV21].

## Definition (Symbolic automaton)

A **nondeterministic symbolic automaton (over $\sigma$)** is a tuple $(\Phi, Q, q_0, F, \delta)$ where:

- $\Phi$ is a finite set of first-order formulas from $\mathcal{F}_\sigma$ called **labels**
- $Q$ is a finite set of **states**,
- $q_0 \in Q$ is the **initial** state,
- $F \subseteq Q$ is the set of **accepting** states, and
- $\delta : Q \times Q \to \Phi$ is the **transition** function.

# Symbolic automata

## Definition (Symbolic automaton — intuition)

A **symbolic automaton** is an automaton with formulae on transitions. Actually, there is much more to them [DV21].

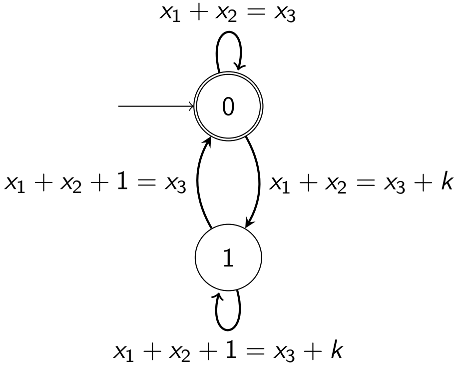
## Definition (Symbolic automaton)

A **nondeterministic symbolic automaton (over $\sigma$)** is a tuple $(\Phi, Q, q_0, F, \delta)$ where:

- $\Phi$ is a finite set of first-order formulas from $\mathcal{F}_\sigma$ called **labels**
- $Q$ is a finite set of **states**,
- $q_0 \in Q$ is the **initial** state,
- $F \subseteq Q$ is the set of **accepting** states, and
- $\delta : Q \times Q \to \Phi$ is the **transition** function.

To read $m$-tuples of natural numbers, we will take $\sigma = \mathbb{N}^m$ and $\mathcal{F}_\sigma$ to be quantifier-free formulae of $(\mathbb{N}; +, |_2, |_3, \dots, k)$ with free variables $x_1, \dots, x_m$, representing the digits in the currently read symbol.

# A symbolic automaton




This automaton recognizes the set $\{(x_1, x_2, x_3) \mid x_1 + x_2 = x_3\}$ in every base $k$.
The state stores the carry.

# Parametric Büchi-Bruyère theorem

**Lemma**

*Symbolic automata with LIA labels can be determinized, and are closed under language union, negation and projection.*

# Parametric Büchi-Bruyère theorem

## Lemma

*Symbolic automata with LIA labels can be determinized, and are closed under language union, negation and projection.*

## Theorem ([KMS26])

*Consider a first-order PBA formula with $m$ free variables and a symbolic automaton recognizing a language of $m$-tuples closed under appending zeroes. We say they are equivalent if for each $k$, the language defined by the symbolic automaton is exactly the $\mathsf{lsdf}_k$ encodings of the set defined by the formula when interpreted in $BA_k$.*
*Such pairing always exists and is effective.*

# Parametric Büchi-Bruyère theorem

**Lemma**

*Symbolic automata with LIA labels can be determinized, and are closed under language union, negation and projection.*

**Theorem ([KMS26])**

*Consider a first-order PBA formula with $m$ free variables and a symbolic automaton recognizing a language of $m$-tuples closed under appending zeroes. We say they are equivalent if for each $k$, the language defined by the symbolic automaton is exactly the $\mathsf{lsdf}_k$ encodings of the set defined by the formula when interpreted in $BA_k$.*
*Such pairing always exists and is effective.*

**Proof.**

From formula to automaton: construct automaton by structural induction.
From automaton to formula: an adaptation of the proof in [BHMV94]. □

# Conclusions

## Corollary

*The set of bases in which a first-order formula of parametric BA is true is effectively definable in LIA.*

*The computation cannot be done essentially faster than TOWER.*

# Conclusions

## Corollary

*The set of bases in which a first-order formula of parametric BA is true is effectively definable in LIA.*

*The computation cannot be done essentially faster than TOWER.*

## Theorem ([KMS26])

*Every first-order PBA formula is equivalent to a $\Sigma_3$ formula.*

*There exists a first-order formula with one free variable that is not equivalent to any $\Sigma_1$ formula.*

# Conclusions

**Corollary**

*The set of bases in which a first-order formula of parametric BA is true is effectively definable in LIA.*

*The computation cannot be done essentially faster than TOWER.*

**Theorem ([KMS26])**

*Every first-order PBA formula is equivalent to a $\Sigma_3$ formula.*

*There exists a first-order formula with one free variable that is not equivalent to any $\Sigma_1$ formula.*

**Mind the gap!**

We don't know whether $\Sigma_2$ is enough.

# Other results

- Augmenting PBA with multiplication by base $k$ (resulting in $(\mathbb{N}; +, V_k, x \mapsto k \cdot x)$) leads to undecidability for one alternation.

# Other results

- Augmenting PBA with multiplication by base $k$ (resulting in $(\mathbb{N}; +, V_k, x \mapsto k \cdot x)$) leads to undecidability for one alternation.
- Digitwise minimum $\&_k$ allows to avoid universal quantifiers.

# Other results

- Augmenting PBA with multiplication by base $k$ (resulting in $(\mathbb{N}; +, V_k, x \mapsto k \cdot x)$) leads to undecidability for one alternation.
- Digitwise minimum $\&_k$ allows to avoid universal quantifiers.
- Symbolic automata accelerate arithmetic solving in all but very small bases.

# Further research

- Close the gap for PBA collapse.

# Further research

- Close the gap for PBA collapse.
- Parametric $(a, b)$-Zeckendorf systems (we have a partial picture).

# Further research

- Close the gap for PBA collapse.
- Parametric $(a, b)$-Zeckendorf systems (we have a partial picture).
- Results for uncountable fields like $\mathbb{R}$ or $\mathbb{Q}_p$ by $\omega$-automata.

# Further research

- Close the gap for PBA collapse.
- Parametric $(a, b)$-Zeckendorf systems (we have a partial picture).
- Results for uncountable fields like $\mathbb{R}$ or $\mathbb{Q}_p$ by $\omega$-automata.

**Thank you!**

# References I


Véronique Bruyère, Georges Hansel, Christian Michaux, and Roger Villemaire.
Logic and *p*-recognizable sets of integers.
*Bulletin of the Belgian Mathematical Society - Simon Stevin*, 1, 01 1994.

Bernard Boigelot.
Symbolic methods and automata.
In *Handbook of Automata Theory*. 2021.

J. Richard Büchi.
Weak second-order arithmetic and finite automata.
*Mathematical Logic Quarterly*, 6(1-6):66–92, 1960.

Alan Cobham.
On the base-dependence of sets of numbers recognizable by finite automata.
*Mathematical systems theory*, 3(2):186–192, Jun 1969.

Loris D'Antoni and Margus Veanes.
Automata modulo theories.
*Communications of the ACM*, 2021.

# References II


Seymour Ginsburg and Edwin H. Spanier.
Semigroups, presburger formulas, and languages.
*Pacific Journal of Mathematics*, 16(2):285–296, 1966.

Kurt Gödel.
Über formal unentscheidbare sätze der principia mathematica und verwandter systeme i.
*Monatshefte für Mathematik und Physik*, 38(1):173–198, 1931.

Christoph Haase and Jakub Różycki.
On the expressiveness of büchi arithmetic.
In Stefan Kiefer and Christine Tasson, editors, *Foundations of Software Science and Computation Structures*, pages 310–323, Cham, 2021. Springer International Publishing.

Jakub Kaszycki, Alessio Mansutti, and Mikhail Starchak.
Parametric recognizability over the integers.
Unpublished work, March, April 2026.

# References III


Mojżesz Presburger.
Über die vollständigkeit eines gewissen systems der arithmetik ganzer zahlen, in welchem die addition als einzige operation hervortritt.
In *Sprawozdanie z I Kongresu Matematyków Krajów Słowiańskich, Warszawa 1929*, pages 92–101, Warsaw, 1929.