# Spectra and satisfiability for logics with successor and a unary func- tion

Arthur Milchior

# Spectra and satisfiability for logics with successor and a unary function

Arthur Milchior *

**Abstract**

We investigate the expressive power of two logics, both with the successor function: first-order logic with an uninterpreted function, and existential monadic second order logic — that is first-order logic over words —, with multiplication by a constant $b$. We prove that all $b$-recognizable sets are spectra of those logics. Furthermore, it is proven that some encoding of the set of halting times of a non-deterministic 2-counter automaton is also a spectrum. This yields undecidability of the finite satisfiability problem for those logics. Finally, it is shown that first-order logic with one uninterpreted function and successor can encode quickly increasing functions, such as the Knuth's up-arrows.

## 1 Introduction

This paper considers the expressive power of fragments of logics with the successor function and a unary (un)interpreted function.

We first recall related results about logics with uninterpreted functions. The classical decision problem [1] consists in giving an algorithm which, for a given logical formula, decides if it is satisfiable. This problem is known to be undecidable in full generality [2]. More precisely, even the satisfiability of first-order logic with two uninterpreted unary functions, which is denoted by $\mathsf{FO}[f(x), g(x)]$, is undecidable [3]. On the other hand, some special cases are decidable, such as first-order logic with one uninterpreted unary function $\mathsf{FO}[f(x)]$ [4].

A formula is finitely satisfiable if and only if admits a finite model. More generally, we may want to study the set of finite models of a formula. A natural approach to understand this set is to consider the *spectrum* of a formula, which is the set of cardinalities of finite models. In particular for a logic over words, the spectrum is the set of size of words defined by a formula. The study of spectra of a logical fragment is a classical topic of descriptive complexity theory. It is usually hard to describe the set of spectra of a logic. The reader is referred to the survey [5]. The class of spectra of $\mathsf{FO}[f(x)]$ — with $f(x)$ uninterpreted — is the class of ultimately periodic sets [6]. The result is also true for $MSO[f(x)]$, that is, Monadic Second-Order logic with one uninterpreted function, [7].

The preceding logics were interpreted over arbitrary universe. Let us now mention some results about logics with some interpreted arithmetic functions only. It is well known that, over integers, $\mathsf{FO}[+]$ is decidable [8], while $\mathsf{FO}[+, \times]$ is undecidable [2].

The class of spectra of $\mathsf{FO}[+1]$ is the class of finite or co-finite sets [9, Theorem IV.3.3.][1] and the class of spectra of $\mathsf{MSO}[+1]$ is the class of ultimately periodic sets [9, Theorem III.2.1].

In this paper, we study logics which mix uninterpreted functions with interpreted (weak) arithmetic functions. Those logics arise in verification problems (see e.g. [10]) and also in descriptive complexity theory (see e.g. [11]). On the one hand, satisfiability of $\Sigma_1[+, <, (f_{i,j}\,(x_1, \ldots, x_j))_{i,j\in\mathbb{N}}]$ is decidable [12]. On the other hand, $\mathsf{MSO}[+1, f(x)]$, is undecidable over finite models and also over $\mathbb{N}$ [13].

*Corresponding author E-mail: Arthur.Milchior@liafa.univ-paris-diderot.fr

[1] This assertion is the special case of the Theorem for a language with a single letter.

In this paper, it is shown that adding the successor function to the logic of an (un)interpreted unary function highly increases the expressive power of this fragment, even when it is limited to the universal-existential first-order fragment.

The case of $\exists\mathsf{MSO}[+1, \times b]$ for $b \geq 2$ is studied in Section 3. It is proven that the set of spectra of the logic $\exists\mathsf{MSO}\Pi_2[+1, \times b]$, that is $\Pi_2[+1, \times b]$ over words, strictly contains the $b$-recognizable sets [9], that is, the sets of integers accepted by an automaton reading base-$b$ digits. Note that by Cobham's theorem ([14], see [15]), a set is $b$-recognizable for all $b \geq 2$ if and only if it is ultimately periodic, hence if and only if it is an $\mathsf{FO}[f(x)]$-spectrum.

We also prove that the class of $\exists\mathsf{MSO}\Pi_2[+1, \times b]$-spectra also contains an encoding of the set of halting times of a non-deterministic 2-counter automaton, a Turing-complete model. It implies that the satisfiability of this logic is undecidable. This result is a generalization of an undecidability result of [16]. Moreover we show that the undecidability result and the encoding of the non-deterministic 2-counter automata result still hold when $\times b$ is replaced by an increasing function whose image is co-infinite. More precisely, it is shown that those results also hold with the restriction that at most two distinct first-order variables can be quantified.

Then, those results are used to study the case $\mathsf{FO}[+1, f(x)]$ with $f$ uninterpreted, in Section 4. It is proven that the class of $\mathsf{FO}[+1, f(x)]$-spectra strictly contains the boolean combination of $b$-recognizable sets, for different $b \geq 2$. It also contains an encoding of the set of halting times of a non-deterministic 2-counter automaton. This, again, leads to the undecidability of satisfiability for this logic.

Another way to assert the expressivity of a logic is to exhibit some sets that it can define. It is proven that $\Pi_2[+1, f(x)]$ — with $f(x)$ uninterpreted — can encode in $\mathbb{N}$ some functions which are increasing extremely quickly such as $n \mapsto c \uparrow^d n$, the Knuth's up-arrow notation, as defined in [17] .

# 2 Definitions

The definitions, notations and useful results used in this paper are introduced in this section.

Let $\mathbb{N}$ be the set of non-negative integers and let $\mathbb{N}^{>0}$ be the set of positive integers. For $n, p \in \mathbb{N}$, let $[n]$ denote $\{i \in \mathbb{N} \mid i \leq n\}$ and $[n, p]$ denote $\{i \in \mathbb{N} \mid n \leq i \leq p\}$. A set $S$ is *co-finite* if $\mathbb{N} \setminus S$ is finite. Let $\#S$ denote the cardinality of $S$.

Let $\mathbb{S}$ and $\mathbb{T}$ be two sets, and let $f : \mathbb{S} \to \mathbb{T}$. Then for $S \subseteq \mathbb{S}$, let $f(S)$ denote $\{f(s) \mid s \in S\}$. For example, if $f$ is $+1$ and $S \subseteq \mathbb{N}$ is the set of even numbers, then $S + 1$ is the set of odd numbers.

## 2.1 Logic

Our definition of the logical formalism is not as general as possible (see e.g. [18]), but corresponds precisely to the notions of logic that are needed to formalize our proofs. First let us define the universe and the vocabularies.

**Definition 1** (Universe)**.** An *universe* $\mathcal{U}$ is a non-empty set. In this paper, the universe is either $\mathbb{N}$, or $[n]$ for some $n \in \mathbb{N}$.

**Definition 2** (Vocabulary)**.** A *vocabulary* is a set with $n + p$ elements

$$\mathcal{V} = \{R_0/d_0, \ldots, R_{n-1}/d_{n-1}, c_0, \ldots, c_{p-1}\}.$$

The $R_i$'s are the *relation* symbols and their arity is $d_i$ and the $c_i$'s are the *constant* symbols.

Note that in this paper, only unary and binary relations are used.

Then let us define the $\mathcal{V}$-structures.

**Definition 3** (Structures)**.** Let $\mathcal{V}$ be a vocabulary. A *$\mathcal{V}$-structure* $\mathcal{S}$ over an universe $\mathcal{U}$ – which is either $\mathbb{N}$, or $[n]$ for some $n \in \mathbb{N}$ – is a tuple:

$$\left(\mathcal{U}, R_0^{\mathcal{S}}, \ldots R_{n-1}^{\mathcal{S}}, c_0^{\mathcal{S}}, \ldots, c_{p-1}^{\mathcal{S}}\right),$$

where $R_i^{\mathcal{S}} \subseteq \mathcal{U}^{d_i}$ and $c_i^{\mathcal{S}} \in \mathcal{U}$.

Given a $\mathcal{V}$-structure $\mathcal{S}$ over $\mathbb{N}$ and an integer $n \in \mathbb{N}^{>0}$ greater than all constants of $\mathcal{S}$, let $\mathcal{S}_{|n}$ be the restriction of $\mathcal{S}$ over the universe $[n-1]$, that is, the $\mathcal{V}$-structure such that ${R_i}^{\mathcal{S}_{|n}} = {R_i}^{\mathcal{S}} \cap [n-1]^{d_i}$ and ${c_i}^{\mathcal{S}_{|n}} = {c_i}^{\mathcal{S}}$.

For every relation symbol of arity $d$ (respectively, constant symbol) $\varsigma$, for every $s \subseteq \mathcal{U}^d$ (respectively, $s \in \mathcal{U}$), let $\mathcal{S}[\varsigma/s]$ be the $\mathcal{V} \cup \{\varsigma\}$-structure with universe $\mathcal{U}$ where $\varsigma^{\mathcal{S}[\varsigma/s]} = s$, and $\tau^{\mathcal{S}[\varsigma/s]} = \tau^{\mathcal{S}}$ for every symbol $\tau$ different from $\varsigma$.

When a symbol $\varsigma$ has a standard interpretation in $\mathbb{N}$ such as an integer, the addition of a constant or the multiplication by a constant, only structures such that $\varsigma^{\mathcal{S}}$ has its usual interpretation are considered. For example "+1" is the successor function, and "$\times b$" is the function that multiplies by $b$.

Furthermore, if $\mathcal{V}$ contains only symbols with a standard interpretation then we simply denote by $\mathbb{N}$ (respectively, $[n]$) the $\mathcal{V}$-structure over universe $\mathbb{N}$ (respectively, $[n]$).

The *fragments* of first- and second-order logic used in this paper are now defined.

**Definition 4** ($\mathcal{V}$-Formulas)**.** Let $\mathcal{V}$ be a vocabulary. The set of quantifier-free $\mathcal{V}$-formulas, denoted by $\Sigma_0[\mathcal{V}]$, is defined by the grammar:

$$\Sigma_0[\mathcal{V}] ::= \neg\varphi_0 \mid \varphi_0 \wedge \varphi_1 \mid \varphi_0 \vee \varphi_1 \mid R_i(c_0, \ldots, c_{d_i-1}) \mid c_0 \doteq c_1$$

where the $c_i$'s are constant symbols and the $\varphi_i$'s are $\Sigma_0[\mathcal{V}]$-formulas.

The sets $\Sigma_i[\mathcal{V}]$ and $\Pi_i[\mathcal{V}]$ are defined by mutual recursion on $i$. For $i \in \mathbb{N}^{>0}$, let $\Sigma_i[\mathcal{V}]$ be defined by the grammar:

$$\Sigma_i[\mathcal{V}] ::= \exists x.\psi \mid \varphi_0 \wedge \varphi_1 \mid \varphi_0 \vee \varphi_1 \mid \neg\chi \mid \xi$$

where $\psi$ is a $\Sigma_i[\mathcal{V}, x]$-formula, the $\varphi_i$'s are $\Sigma_i[\mathcal{V}]$-formulas, $\chi$ is a $\Pi_i[\mathcal{V}]$-formula and $\xi$ is either a $\Pi_{i-1}[\mathcal{V}]$-formula or a $\Sigma_{i-1}[\mathcal{V}]$-formula. Similarly, let $\Pi_i[\mathcal{V}]$ be defined by the grammar:

$$\Pi_i[\mathcal{V}] ::= \forall x.\psi \mid \varphi_0 \wedge \varphi_1 \mid \varphi_0 \vee \varphi_1 \mid \neg\chi \mid \xi$$

where $\psi$ is a $\Pi_i[\mathcal{V}, x]$-formula, the $\varphi_i$'s are $\Pi_i[\mathcal{V}]$-formulas, $\chi$ is a $\Sigma_i[\mathcal{V}]$-formula and $\xi$ is either a $\Pi_{i-1}[\mathcal{V}]$-formula or a $\Sigma_{i-1}[\mathcal{V}]$-formula.

For a logical fragment $\mathcal{L}[\mathcal{V}]$, let $\exists\mathsf{MSO}\mathcal{L}[\mathcal{V}]$ (respectively, $\forall\mathsf{MSO}\mathcal{L}[\mathcal{V}]$) denote the set of formulas of the form $\exists R_0, \ldots, R_n.\psi$ (respectively, $\forall R_0, \ldots, R_n.\psi$) where $\psi$ is a $\mathcal{L}[\mathcal{V}, R_0, \ldots, R_n]$-formula, and the arity of the $R_i$'s is 1.

For the sake of clarity, we state that $\varphi(x_0, \ldots, x_{d-1})$ is a $\mathcal{L}[\mathcal{V}]$-formula when $\varphi$ is a $\mathcal{L}[\mathcal{V} \cup \{x_0, \ldots, x_{d-1}\}]$-formula.

Section 3 considers results which concern two-variable logic, which is now introduced.

**Definition 5** (Two-variable logic)**.** Let us assume that the vocabulary $\mathcal{V}$ contains at most two constant symbols $x$ and $y$. The set of two-variable $\mathcal{L}[\mathcal{V}]$-formulas, denoted $\mathcal{L}^2[\mathcal{V}]$, is the set of formulas where the only quantified first-order variables are $x$ and $y$.

A formula is said to be in prenex-normal form if it contains a sequence of quantifications, and then a quantifier-free formula. The set of formulas $\Sigma_i[\mathcal{V}]$ and $\Pi_i[\mathcal{V}]$ are usually defined as the set of formulas in prenex-normal form formulas with $i-1$ alternations. Let us call it the PNF definition. The PNF definition and our definition are equivalent in terms of definability. On the other hand, there are two-variable formulas which are not equivalent to two-variable formulas in prenex-normal form. Hence, in terms of PNF, what we call $\Pi_i^2[\mathcal{V}]$ can be defined as "the set of $\mathsf{FO}^2[\mathcal{V}]$-formulas equivalent to a $\Pi_i[\mathcal{V}]$-formula".

This notation allows to consider simultaneously the number of variables and the number of alternation of quantifiers in our formulas.

A notation for implication and equivalence is now introduced.

**Notation 6.** Let $\mathcal{V}$ be a vocabulary. Let $\varphi \in \Sigma_i[\mathcal{V}]$ and $\psi \in \Pi_i[\mathcal{V}]$, then let $\varphi \Longrightarrow \psi$ be syntactical sugar for the $\Pi_i[\mathcal{V}]$-formula $\neg\varphi \vee \psi$. Then let $\varphi \Longleftrightarrow \psi$ be syntactical sugar for the formula $(\varphi \Longrightarrow \psi) \wedge (\psi \Longrightarrow \varphi)$ which belongs simultaneously to $\Pi_{i+1}[\mathcal{V}]$ and to $\Sigma_{i+1}[\mathcal{V}]$.

If $\varphi$ and $\psi$ are two-variable formulas, then $\varphi \Longrightarrow \psi$ and $\varphi \Longleftrightarrow \psi$ are also two-variable formulas.

**Definition 7** (Interpretation and Model)**.** Let $\mathcal{V}$ be a vocabulary, $\mathcal{S}$ be a $\mathcal{V}$-structure over an universe $\mathcal{U}$ as in Definition 1. That is, $\mathcal{U}$ is either $\mathbb{N}$ or $[n]$ for some $n \in \mathbb{N}$.

Let $\varphi$ be a $\mathcal{V}$-formula. We define the satisfaction relation $\mathcal{S} \models \varphi$ by induction over $\varphi$ as follows:

- $\mathcal{S} \models c \doteq c'$ if and only if $c^{\mathcal{S}} = c'^{\mathcal{S}}$,
- $\mathcal{S} \models R(t)$ if and only if $t^{\mathcal{S}} \in R^{\mathcal{S}}$,
- $\mathcal{S} \models \psi_0 \vee \psi_1$ if and only if $\mathcal{S} \models \psi_0$ or $\mathcal{S} \models \psi_1$,
- $\mathcal{S} \models \psi_0 \wedge \psi_1$ if and only if $\mathcal{S} \models \psi_0$ and $\mathcal{S} \models \psi_1$,
- $\mathcal{S} \models \neg\psi$ if and only if $\mathcal{S} \not\models \psi$,
- $\mathcal{S} \models \exists x.\psi$ if and only if there exists $n \in \mathcal{U}$ such that $\mathcal{S}[x/n] \models \psi$,
- $\mathcal{S} \models \forall x.\psi$ if and only if for all $n \in \mathcal{U}$, $\mathcal{S}[x/n] \models \psi$,
- $\mathcal{S} \models \exists R.\psi$ if and only if there exists $N \subseteq \mathcal{U}$ such that $\mathcal{S}[R/N] \models \psi$,
- $\mathcal{S} \models \forall R.\psi$ if and only if for all $N \subseteq \mathcal{U}$, $\mathcal{S}[R/N] \models \psi$.

A $\mathcal{V}$-structure $\mathcal{S}$ is said to be a model of $\varphi$ if $\mathcal{S} \models \varphi$.

Sometimes we need to restrict quantifiers to elements that satisfy a certain property $\chi$.

**Notation 8** (Relativization)**.** If $\chi$ is a formula, we write $(\forall x.\chi)\,\varphi$ for $\forall x.\,(\chi \Longrightarrow \varphi)$ and $(\exists x.\chi)\,\varphi$ for $\exists x.\,(\chi \wedge \varphi)$.

We now explain how logical formulas define sets.

**Definition 9** (Definability)**.** Let $\varphi(x_0, \ldots, x_{d-1})$ be a formula with $d$ free variables in a logic $\mathcal{L}[\mathcal{V}]$. Given a $\mathcal{V}$-structure $\mathcal{S}$, we say that $\varphi$ defines in $\mathcal{S}$ the $d$-ary set

$$\varphi(x_0, \ldots, x_{d-1})^{\mathcal{S}} = \{(n_0, \ldots, n_{d-1}) \in \mathcal{U}^d \mid \mathcal{S} \models \varphi(n_0, \ldots, n_{d-1})\}.$$

We say that a set $R \subseteq \mathcal{U}^d$ is $\mathcal{L}[\mathcal{V}]$-definable in $\mathcal{S}$ if there exists $\varphi(x_0, \ldots, x_{d-1}) \in \mathcal{L}[\mathcal{V}]$ such that $R$ is equal to $\varphi(x_0, \ldots, x_{d-1})^{\mathcal{S}}$. Furthermore, if $f$ is a function from $\mathcal{U}$ to $\mathcal{U}$, we say that $f$ is $\mathcal{L}[\mathcal{V}]$-definable if its graph $\{(x, y) \in \mathcal{U}^2 \mid f(x) = y\}$ is $\mathcal{L}[\mathcal{V}]$-definable.

**Functions** Let us say a word about functions in this paper. Usually in finite model theory and in two-variable logic, the vocabulary does not contain function symbols, hence the only terms are the constants. Instead, formally, unary functions $f$ are replaced with a binary relation $f(x, y)$, such that for every $n \in \mathcal{U}$ there is at most one $r \in \mathcal{U}$ such that $f(n, r)$ holds. For example, the addition of 1 is denoted as the binary relation $+1(x, y)$ interpreted in $[n]$ by $\{(x, x+1) \mid x < n\}$. The distinction is important because over the universe $[n]$, the value of $n + 1$ is undefined.

On the other hand, for the sake of clarity, "$+1(x, y)$" is written "$x + 1 \doteq y$", and more generally if $f$ is a unary function, then "$f(x, y)$" is written "$f(x) \doteq y$".

In particular, every vocabulary considered in this paper contains the successor function. Hence, the notation $x + c$ for $c \in \mathbb{N}$ is used, as an abbreviation for $x + (1 + (\ldots (1 + 1) \ldots))$.

In this paper, we specify the number of alternations of quantifiers and the number of variables used in formulas. For the sake of simplicity, we introduce abbreviations such as $f(g(x))$. The following lemmas explain how to encode formally all of those abbreviations in fragments of the logic.

**Lemma 10.** *Let $f$ and $g$ be two unary function symbols, then $f(g(x)) = y$ is equivalent to a $\Sigma_1[f, g]$-formula.*

*Proof.* The formula $f(g(x)) = y$ is equivalent to $\exists z.g(x) \doteq z \wedge f(z) \doteq y$ which belongs to $\Sigma_1[f, g]$. Clearly, if $g^{\mathcal{S}}(x^{\mathcal{S}})$ is not defined then this formula does not hold in $\mathcal{S}$, which is the intended behaviour as $f(g(x))$ is not defined either. □

For $R$ a unary relation symbol, the notation $R(f(g(x)))$ can also be used in two-variable logic.

**Lemma 11.** *Let $f$ and $g$ be two unary function symbols, and let $R$ be a unary relation symbol, then $R(f(g(x)))$ is equivalent to a $\Sigma_1^2[f, g, x]$-formula.*

*Proof.* The formula $R(f(g(x)))$ is equivalent to

$$\exists y.g(x) \doteq y \wedge \{\exists x.f(y) \doteq x \wedge R(x)\}.$$

□

The two preceding lemmas extend easily to the application of more than two function symbols.
In general, $f(g(x)) \doteq y$ is not equivalent to a $\Sigma_1^2[f, x, y]$-formula. Indeed, a formula defining the set

$$\{(x, y) \mid f(g(x)) = y\}$$

should quantify a variable to store the value of $g(x)$, but in two-variable logic, there is no guarantee that such a variable is available.

**Constants** Similarly, for the sake of simplicity, our formula uses constant symbols, such as 2 or $(\text{last} -2)$ where the value "last" corresponds to the greatest integer of the set.

**Lemma 12.** *The formulas $x \doteq c$ and $x \doteq \text{last} - c$ for $c \in \mathbb{N}$, are equivalent to a $\Sigma_2^2[+1, x]$-formula and to a $\Pi_2^2[+1, x]$-formula.*

*Proof.* By symmetry, the proofs are similar for $x \doteq c$ and $x \doteq \text{last} - c$. Therefore, we only prove the first case. The proof proceeds by induction over $c$. If $c = 0$ then $x \doteq 0$ is equivalent to $\forall y.y + 1 \neq x$. If $c \in \mathbb{N}^{>0}$, then $x \doteq c$ is equivalent to

$$\exists y.y + 1 \doteq x \wedge y \doteq (c - 1)$$

and to

$$\{\exists y.y + 1 \doteq x\} \wedge \{\forall y.y + 1 \doteq x \Longrightarrow y \doteq c - 1\}.$$

□

This result is used to use constants as parameters of unary relations as explained in the following lemma.

**Lemma 13.** *The formulas $R(c)$ and $R(\text{last} - c)$, for $c \in \mathbb{N}$, are equivalent to a $\Sigma_2^2[+1]$-formula and to a $\Pi_2^2[+1]$-formula.*

*Proof.* As in the previous proof, we consider only the case $R(c)$ since the other one is similar. The formula $R(c)$ is equivalent to $\exists x.x \doteq c \wedge R(x)$ and to $\forall x.x \doteq c \Longrightarrow R(x)$. □

## 2.2 Satisfiability

Let $\varphi$ be a $\mathcal{L}[\mathcal{V}]$-formula without free variable. Then we say that $\varphi$ is *(finitely) satisfiable* if there exists a (finite) $\mathcal{V}$-structure $\mathcal{S}$ such that $\mathcal{S} \models \varphi$.

The next lemma extends undecidability of the finite satisfiability problem to undecidability over $\mathbb{N}$. This will allow to prove theorems about finite universes only.

**Lemma 14.** *Let $\mathcal{L}[\mathcal{V}]$ be a first-order logic with universal and existential quantifiers, and at least two first-order variables, such that finite satisfiability for $\exists\mathsf{MSO}\mathcal{L}[\mathcal{V}]$ is undecidable. Then satisfiability of $\exists\mathsf{MSO}\mathcal{L}[\mathcal{V} \cup \{+1\}]$ is undecidable over $\mathbb{N}$.*

*Proof.* Let $\varphi \in \mathcal{L}[\mathcal{V}]$ and $\mathcal{S}$ be a $\mathcal{V}-$structure of cardinality $n$. Let $M$ be a second-order variable not in $\mathcal{V}$, and let $\mathcal{S}'$ be the $(\mathcal{V} \cup \{M\})$-structure over $\mathbb{N}$, where $M^{\mathcal{S}'} = [n-1]$, $+1$ has its usual interpretation, and $\varsigma^{\mathcal{S}} = \varsigma^{\mathcal{S}'}$ for every symbol $\varsigma$ of $\mathcal{V}$.

Let $\varphi'$ be obtained from $\varphi$ by replacing each quantification $\exists x.\psi$ by $\exists x.\,(M(x) \wedge \psi)$ and $\forall x.\psi$ by $\forall x.\,(M(x) \Longrightarrow \psi)$. By an easy induction over $\varphi$, we have $\mathcal{S} \models \varphi \Longleftrightarrow \mathcal{S}' \models \varphi'$, and in particular, $\varphi$ has a finite model if and only if the formula

$$\exists M.\,\{[\forall x \neq 0.M(x) \Longrightarrow M(x-1)] \wedge [\exists y.\neg M(y)] \wedge M(0) \wedge \varphi'\}$$

has a model of universe $\mathbb{N}$. □

## 2.3 Spectra

The main notion of this paper is now introduced.

**Definition 15** (Spectra)**.** Let $\varphi$ be a formula without free variable. Its spectrum $\mathsf{SP}(\varphi)$ is the set of positive numbers $n$ such that $\varphi$ has a model of cardinality $n$. If $\mathcal{F}$ is a set of formulas then $\mathsf{SP}(\mathcal{F})$ is defined as the set of spectra of formulas in the set $\mathcal{F}$.

A set $S \subseteq \mathbb{N}^{>0}$ is said to be a $\mathcal{L}[\mathcal{V}]$-spectrum if there exists a $\mathcal{L}[\mathcal{V}]$-formula $\varphi$ such that $\mathsf{SP}(\varphi) = S$.

Since the empty universe is not considered, the integer 0 does not belong to any spectrum. The following lemma should be noted:

**Lemma 16.** *Let $\psi$ be a formula, then $\mathsf{SP}(\exists x.\psi) = \mathsf{SP}(\psi)$ and $\mathsf{SP}(\exists R.\psi) = \mathsf{SP}(\psi)$.*

Three lemmas are given in this section. They help to create $\exists\mathsf{MSO}[+1, \mathcal{V}]$-spectra.

The following lemma states that the set of spectra is closed by adding or removing finitely many integers.

**Lemma 17.** *Let $\mathcal{V}$ be a vocabulary which contains $+1$ and $\mathcal{L}[\mathcal{V}]$ be a logical fragment that contains $\Pi_2^2[\mathcal{V}]$. Let $S$ be a spectrum of a $\mathcal{L}[\mathcal{V}]$-formula. Let $F$ and $F'$ be finite disjoint subsets of $\mathbb{N}^{>0}$. Then $(S \cup F) \setminus F'$ is a $\mathcal{L}[\mathcal{V}]$-spectrum.*

*Proof.* Assume that there exists formulas $\gamma_{n,\Pi} \in \Pi_2^2[+1]$ and $\gamma_{n,\Sigma} \in \Sigma_2^2[+1]$, which hold only on models of cardinality $n$. Let $\varphi_S$ be the formula whose spectrum is $S$, then $(S \cup F) \setminus F'$ is the spectrum of:

$$\left(\varphi_S \vee \bigvee_{n \in F} \gamma_{n,\Pi}\right) \wedge \bigwedge_{n \in F'} \neg\gamma_{n,\Sigma}.$$

The formulas $\gamma_{n,\Sigma}$ and $\gamma_{n,\Pi}$ must now be defined. Assume that there exist $\gamma_{\geq n,\Pi} \in \Pi_2[+1]$ and $\gamma_{\geq n,\Sigma} \in \Sigma_2[+1]$, which state that universe's cardinality is at least $n$. Then $\gamma_{n,\Pi}$ can be defined as $\gamma_{\geq n,\Pi} \wedge \neg\gamma_{\geq n+1,\Sigma}$ and $\gamma_{n,\Sigma}$ can be defined as $\gamma_{\geq n,\Sigma} \wedge \neg\gamma_{\geq n+1,\Pi}$.

The formulas $\gamma_{\geq n,\Sigma}$ and $\gamma_{\geq n,\Pi}$ must now be defined. Let $\gamma_{\geq n,\Sigma}$ be $\exists x.x \doteq n-1$. Assume that there exists a $\Pi_2^2[+1, x]$ formula $suc_n(x)$ which asserts that $x + n \leq \text{last}$. Then let $\gamma_{\geq n,\Pi}$ be $\forall x.x \doteq 0 \Longrightarrow suc_{n-1}(x)$.

Now we define the formula $suc_n(x)$ by induction over $n$. This formula is similar to the formula of the proof of Lemma 12. For $n = 0$, we can define $suc_0(x)$ as the formula "$\mathrm{true}$". For $n > 0$ let $suc_n(x)$ be

$$\{\exists y.x + 1 \doteq y\} \wedge \{\forall y.x + 1 \doteq y \wedge suc_{n-1}(y)\}.$$

□

The following lemma states that spectra are closed by union.

**Lemma 18.** *Let $\mathcal{V}$ be a vocabulary and $\mathcal{L}$ a logic. Let $(S_i)_{i \in [d-1]}$ be a sequence of $d$ $\mathcal{L}[\mathcal{V}]$-spectra. Then $S = \bigcup_{i=0}^{d-1} S_i$ is a $\mathcal{L}[\mathcal{V}]$-spectrum.*

*Proof.* Let $\varphi_i \in \mathcal{L}[\mathcal{V}]$ be such that $\mathsf{SP}(\varphi_i) = S_i$. Let $\varphi = \bigvee_{i=0}^{d-1} \varphi_i$. Then $\mathsf{SP}(\varphi) = S$. □

The following lemma proves that for any logic that contains the successor relation and $\Pi_2^2$, spectra are closed by addition of positive constants.

**Lemma 19.** *Let $\mathcal{V}$ be a vocabulary that contains $+1$ and let $\mathcal{L}[\mathcal{V}]$ be a logical fragment that contains $\Pi_2^2[\mathcal{V}]$. Let $S \subseteq \mathbb{N}^{>0}$ be a $\mathcal{L}[\mathcal{V}]$-spectrum. Let $a \in \mathbb{N}$. Then $S + a$ is a $\mathcal{L}[\mathcal{V}]$-spectrum.*

*Proof.* We prove the result for $a = 1$. The general case is proven by applying the base case $a$ times.

The proof transforms a $\mathcal{V}$-structure $\mathcal{S}$ of cardinality $n$ into a $\mathcal{V}$-structure $\mathcal{S}'$ of cardinality $n+1$ by adding an integer at the end of the structure.

We replace every quantification by relativized quantification, as explained in Notation 8, to ensure that the interpretation of quantified variables is restricted to $[\text{last} - 1]$. Then, the last integer of the universe cannot be used.

Let $\varphi \in \mathcal{L}[\mathcal{V}]$, then let $\Phi(\varphi)$ be the $\mathcal{L}[\mathcal{V}]$-formula obtained from $\varphi$ by replacing each subformula of the form $\exists x.\psi$ by

$$\exists x. \{[\exists y. x + 1 \doteq y] \wedge \Phi(\psi)\}.$$

and each subformula of the form $\exists R.\psi$ by

$$\exists R. \{\neg R(\text{last}) \wedge \Phi(\psi)\}.$$

Let $\varphi$ be a $\mathcal{L}[\mathcal{V}]$-formula. By induction over $\varphi$, for each $n \in \mathbb{N}^{>0}$, $n \in \mathsf{SP}(\varphi)$ if and only if $n + 1 \in \mathsf{SP}(\Phi(\varphi))$. Hence $\mathsf{SP}(\Phi(\varphi)) \setminus \{1\} = \mathsf{SP}(\varphi) + 1$. By Lemma 18, $\mathsf{SP}(\varphi) + 1$ is also a $\mathcal{L}[\mathcal{V}]$-spectrum. □

## 2.4 Words and automata

An alphabet $A$ is a finite non-empty set. The set $A^*$ (respectively, $A^+$) stands for the set of finite (respectively, non-empty) sequences of letters of $A$, which are called words. Let $\varepsilon$ be the only word of length 0. For each $a \in A$, $a$ is a word of length 1. We denote by $|w|$ the length of $w$.

We shall consider a logic interpreted over words. A logic over words always include the unary set symbols $P_a$ for all $a \in A$. Let $w = w[0] \dots w[n-1]$ with $w[i] \in A$. Let $\mathcal{S}_w$ be the $\{+1, (P_a)_{a \in A}\}$-structure:

$$\left([n-1], \{(i, i+1) \mid i \in [n-2]\}, (\{i \in [n-1] \mid w[i] = a\})_{a \in A}\right).$$

So $P_a(i)$ is true if and only if the $i$-th letter is $a$. More generally, let $\mathcal{V}$ be a vocabulary without the $P_a$ symbols and let $\mathcal{S}$ be a $\mathcal{V}$-structure over $\mathbb{N}$. Then $\mathcal{S}_w$ is the $\mathcal{V} \cup \left\{(P_a)_{a \in A}\right\}$-structure defined as follows:

$$\mathcal{S}_{|n}[(P_a / \{i \in [|w| - 1] \mid w[i] = a\})_{a \in A}].$$

**Definition 20** (Finite Deterministic Automaton)**.** A finite deterministic automaton $\mathcal{A}$ over alphabet $A$ is a quintuple $(Q, A, \delta, q_0, F)$ where $Q$ is the finite set of states, $q_0 \in Q$ is the initial state, $F \subseteq Q$ is the set of accepting states and $\delta : Q \times A \to Q$ is the transition function.

The domain of the function $\delta$ is extended to $Q \times A^*$ by setting $\delta(q, \varepsilon) = q$ and $\delta(q, wa) = \delta(\delta(q, w), a)$ for $a \in A, w \in A^*$. We say that $w \in A^*$ is accepted by $\mathcal{A}$ if $\delta(q_0, w) \in F$. We denote by $\underline{\mathcal{A}}$ the set of words accepted by $\mathcal{A}$ and we say that $\mathcal{A}$ accepts the set $\underline{\mathcal{A}}$.

## 2.5 The $b$-recognizable sets

In this section, the $b$-recognizable sets are introduced, for $b \geq 2$. We refer to [15] for general results about $b$-recognizable sets.

A word over alphabet $[b-1]$ is said to be a word in base $b$. Let $\overline{\cdot}^{\mathbb{N}}$ be the function from $[b-1]^*$ to $\mathbb{N}$ that sends a word $w$ to the natural number it represents, least-digit first. Formally, it is defined as $\overline{w}^{\mathbb{N}} = \sum_{i=0}^{|w|-1} w[i] b^{|w|-i-1}$, or recursively by $\overline{\varepsilon}^{\mathbb{N}} = 0$, and for $a \in [b-1], w \in [b-1]^*$, as $\overline{aw}^{\mathbb{N}} = a + b \cdot \overline{w}^{\mathbb{N}}$. For example, in base $b = 3$, we have $\overline{122}^{\mathbb{N}} = 25$. For $L \subseteq [b-1]^*$, let $\overline{L}^{\mathbb{N}}$ be $\{\overline{w}^{\mathbb{N}} \mid w \in L\}$.

It should be noted that any integer $n$ has an infinite number of representations in base $b$ due to possible leading 0. More precisely if $|w| \geq |w'|$ then $w' \in w0^*$. The least representation of $n \in \mathbb{N}$ in base $b$ is denoted $\underline{n}_{\mathbb{N}}$. It is defined by induction over $n$ by $\underline{0}_{\mathbb{N}} = \varepsilon$ and for $n > 0$ by $\underline{n}_{\mathbb{N}} = (n \mod b) \cdot \underline{\lfloor \frac{n}{b} \rfloor}_{\mathbb{N}}$.

In this paper, we only consider automata which only accept integer representations without leading zero.

**Definition 21** (Automaton in base $b$)**.** An automaton $\mathcal{A} = (Q, [b-1], \delta, q_0, F)$ is said to be an automaton in base $b$ if $\delta(q, 0) \notin F$ for every $q \in Q$.

Let $\overline{\mathcal{A}}^{\mathbb{N}}$ denote $\overline{(\mathcal{A})}^{\mathbb{N}}$, i.e. the set of integers that have one representation accepted by $\mathcal{A}$.

**Definition 22** ($b$-recognizable sets, $b-$ REC)**.** A set $S \subseteq \mathbb{N}^{>0}$, is $b$-recognizable if there exists an automaton $\mathcal{A}$ in base $b$ such that $\overline{\mathcal{A}}^{\mathbb{N}} = S$. Let $b-$ REC denote the class of $b$-recognizable sets.

**Example 23.** The set of powers of $b$, and the set of integers which have an odd number of 1 in base $b$, are $b$-recognizable. The ultimately periodic sets, that is, the sets $S$ such that there exists $N, p \in \mathbb{N}$ such that for all $n \geq N$, $n \in S \Longleftrightarrow n + p \in S$, are $b$-recognizable for all $b$.

We assume that $S$ is a subset of $\mathbb{N}^{>0}$, namely that $0 \notin S$, for technical reasons.

The class $b-$ REC admits a logical characterization. Let $V_b : \mathbb{N} \to \mathbb{N}$ be the function that sends $n$ to the greatest power of $b$ dividing $n$, with the convention that $V_b(0) = 1$.

**Theorem 24** (Büchi-Bruyère Theorem, see [15])**.** *The class of subsets of $\mathbb{N}^{>0}$ which are $\mathsf{FO}[+, V_b]$-definable (in $\mathbb{N}$) is equal to $b-$ REC.*

A set $S$ is said to be *-recognizable if it is a boolean combination of $b$-recognizable sets for possibly different values of $b$. Let $*$REC be the class of *-recognizable sets. Observe that in general there is no notion of automata accepting *-recognizable sets. It is only a strict subset of the $\mathsf{FO}[+, (V_b)_{b\in\mathbb{N}}]$-definable sets.

## 2.6 Non-deterministic 2-counter automata

The definition of a *non-deterministic 2-counter automaton* is briefly recalled. The undecidability of the halting problem for non-deterministic 2-counter automata is used to prove the undecidability of the satisfiability problem of the logics studied in this paper.

**Definition 25** (Non-deterministic 2-counter automata)**.** A non-deterministic 2-counter automaton $\mathcal{A}$ consists of a list of instructions. Let $\#\mathcal{A}$ denote the number of instruction of $\mathcal{A}$. The instructions are "incr($h$)", "decr($h$)", "jmp($i, j$)", "jz($h, j$)" with $h \in \{0, 1\}$, $i, j \in [\#\mathcal{A} - 1]$ and "Halt". The $j$-th instruction is written $\mathcal{A}_j$. Without loss of generality we assume that only one Halt instruction appears in the list and that it appears as the last instruction.

Then we explain how those automata compute.

**Definition 26** (Configuration and Computation)**.** Let $\mathcal{A}$ be a 2-counter automaton. A *configuration* of $\mathcal{A}$ is a 3-tuple of integers $(q, n_0, n_1)$ where $q$ is the next instruction of the automaton and $n_j$ is the value of the $j$-th counter.

Let $(\kappa, c_0, c_1)$ be a triplet of lists of the same length $l$, such that for all $i \in [l-1]$, $(\kappa[i], c_0[i], c_1[i])$ is a configuration. This triplet is a *computation of* $\mathcal{A}$ if the first configuration satisfies $c_0[0] = c_1[0] = \kappa[0] = 0$, the last configuration satisfies $\kappa[l-1] = (\#\mathcal{A} - 1)$, and for every $j \in [l-2]$ such that $\mathcal{A}_{\kappa[j]} \neq$ Halt – that is $\kappa[j] < \#\mathcal{A} - 1$ – we have:

$$
\begin{array}{llllll}
\text{if } \mathcal{A}_{\kappa[j]} = incr(i) & \text{then} & c_i[j+1] = c_i[j] + 1, & c_{1-i}[j+1] = c_{1-i}[j], & \text{and} & \kappa[j+1] = \kappa[j] + 1, \\
\text{if } \mathcal{A}_{\kappa[j]} = decr(i) & \text{then} & c_i[j+1] = \max(c_i[j] - 1, 0), & c_{1-i}[j+1] = c_{1-i}[j], & \text{and} & \kappa[j+1] = \kappa[j] + 1, \\
\text{if } \mathcal{A}_{\kappa[j]} = jmp(n, m) & \text{then} & \forall i. c_i[j+1] = c_i[j], & & \text{and} & \kappa[j+1] \in \{n, m\}, \\
\text{if } \mathcal{A}_{\kappa[j]} = jz(i, m) & \text{then} & \forall i. c_i[j+1] = c_i[j], & \text{if } c_i[j] = 0 \text{ then } \kappa[j+1] = m & \text{and otherwise} & \kappa[j+1] = \kappa[j] + 1.
\end{array}
$$

It should be noted that those automata do not have any input.

The notion of spectra is extended to 2 counter automata.

**Definition 27** ($\mathsf{SP}(\mathcal{A})$)**.** Let $\mathcal{A}$ be a non-deterministic 2-counter automaton. Then we define $\mathsf{SP}(\mathcal{A}) \subseteq \mathbb{N}$ as the set of integers $n$ such that $n \in \mathsf{SP}(\mathcal{A})$ if there is a computation of $\mathcal{A}$ with $n$ steps.

We give an example of a non-deterministic 2-counter automaton, whose spectra is not ultimately periodic. This example is useful for the proof of the strict inclusion of $b$-recognizable sets in $\mathsf{SP}\left(\exists\mathsf{MSO}[+1, \times b]\right)$ (see Corollary 39.)

**Example 28.** Let $\mathcal{A}$ be the automaton which implements Algorithm 1. Formally, it is

(0) $incr\,(0)$ (1) $jmp\,(0,2)$ (2) $decr\,(0)$ (3) $incr\,(1)$ (4) $jz\,(0,2)$ (5) $decr\,(1)$ (6) $incr\,(0)$
(7) $jz\,(1,5)$ (8) $decr\,(0)$ (9) $jz\,(0,2)$ (10) $Halt$

This algorithm non-deterministically chooses a positive integer $n$ in the first counter. It transfers the first counter to the second one, then it transfers back the second counter to the first one, and decrements the first counter. It repeats until the first counter equals 0.

Let us prove that $\mathsf{SP}\left(\mathcal{A}\right) = \{7n + 3n^2 + 1 \mid n \in \mathbb{N}^{>0}\}$.

**Algorithm 1:** Example 28

```
   repeat
0  |  incr(0)
   until non-deterministically chooses when to end; //jmp[0,2]
   repeat
   |  repeat
2  |  |  decr(0)
3  |  |  ncr(1)
   |  until c0 = 0; //jz[0,2]
   |  repeat
5  |  |  decr(1)
6  |  |  incr(0)
   |  until c1 = 0; //jz[1,5]
8  |  decr(0)
   until c0 = 0; //jz[0,2]
10 halt
```

We study an halting computation. The loop of Line 1 is executed $n$ times, for an arbitrary $n \in \mathbb{N}^{>0}$. Each execution of the loop costs 2 steps, hence the entire loops costs $2n$ steps. At the ends of the loops, $c_0 = n$ and $c_1 = 0$.

The outer loop of line 9 is executed $n$ times, the loop variant is that at the end of the $i$-th execution, $c_0 = n - i$ and $c_1 = 0$. During the $i$th execution of the outer loop, the first inner loop, at line 4 is repeated $n - i$ times. The loop variant is such that after the $j$th execution, $c_0 = n - i - j$ and $c_1 = j$. Each execution costs 3 steps, hence the first inner loop costs $3j$ steps. The second inner loop is identical to the first one, with $c_0$ and $c_1$ exchanged. Finally the instruction of line 8 decrements the counter 0. Hence the $i$th execution of the outer loop costs $6\,(n - i) + 2$ steps. Finally, the Halt instruction costs one step.

Finally, the number of steps of the execution is:

$$2n + \sum_{i=0}^{n-1} [6\,(n-i) + 2] + 1 = 2n + 6\sum_{i=1}^{n} i + 2n + 1 = 4n + 6\frac{n\,(n+1)}{2} + 1 = 7n + 3n^2 + 1$$

Hence $\mathsf{SP}\left(\mathcal{A}\right) = \left\{7n + 3n^2 + 1 \mid n \in \mathbb{N}^{>0}\right\}$.

# 3 Existential monadic second-order logic with arithmetic

Theorems about existential monadic logic with arithmetic relations are proven in this section. More precisely, we consider the logic $\exists\mathsf{MSO}\Pi_2^2\,[+1, \times b]$ for $b \geq 2$ and $\exists\mathsf{MSO}\Pi_2^2\,[+1, g]$ for some increasing functions $g$.

Section 3.1 deals with spectra which are $b$-recognizable. Then Section 3.2 deals with spectra which encode spectra of some non-deterministic 2-counter automaton.

### 3.1 $b$-recognizable sets are $\exists\mathsf{MSO}[+1, \times b]$-spectra

In this section, we show how to represent $b$-recognizable sets as $\exists\mathsf{MSO}^2[+1, \times b]$-spectra and as $\forall\mathsf{MSO}^2[+1, \times b]$-spectra.

**Theorem 29.** *The following inclusion holds:*

$$b - \mathit{REC} \subseteq \mathsf{SP}\left(\exists\mathsf{MSO}\Pi_2^2[+1, \times b]\right).$$

Actually Corollary 39 states that this inclusion is strict.
To prove this theorem, the following Lemma is needed.

**Lemma 30.** *Let $\mathcal{A}$ be an automaton in base $b$ as in Definition 21. There exists a $\exists\mathsf{MSO}\Pi_2^2[+1, \times b]$-formula $\varphi_{\mathcal{A}}$ which holds on $[n]$ if and only if $n$ is accepted by $\mathcal{A}$.*

Note that $\mathsf{SP}\left(() \varphi_{\mathcal{A}}\right) = \overline{\mathcal{A}}^{\mathbb{N}} + 1$, indeed, the cardinality of $[n]$ is $n + 1$. Let us first give an example to illustrate the lemma.

**Example 31.** Let $\mathcal{A}$ be the automaton of Figure 1. It accepts the language $0^*1$, hence the set of integers $\{2^i \mid i \in \mathbb{N}\}$.

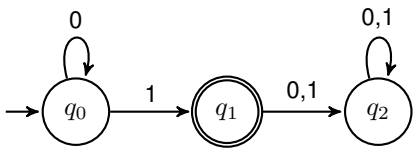


Figure 1: Automaton accepting $\{2^i \mid i \in \mathbb{N}\}$

This set is defined by:

$$\exists R. \forall x. \{R(x) \Longleftrightarrow [x \doteq \mathrm{last} \vee R(2x)]\} \wedge R(1). \tag{1}$$

This formula holds for the structures $\mathcal{S}$ over the universe $[2^c]$ for $c \in \mathbb{N}$ when the variable $R$ is interpreted by $\{2^i \mid i \in [c]\}$. Intuitively, $R$ is interpreted by the set of integers $n$ such that $\delta(q_0, \underline{n}_{\mathbb{N}}) \in F$.

Lemma 30 is now proven.

*Proof.* For $b \in \mathbb{N}$, let $\mathcal{V}_{b,Q}$ be the vocabulary $\{+1, \times b, (R_q)_{q \in Q}\}$ where each $R_q$ is a unary relation for any state $q \in Q$. Let $n \in \overline{\mathcal{A}}^{\mathbb{N}}$. We defined a $\mathcal{V}_{b,Q}$-structure $\mathcal{S}_n$ which encodes the computation of $\mathcal{A}$ on the input $w = \underline{n}_{\mathbb{N}}$ of length $l$. The main idea of the proof is to choose a sequence of integers $c_{n,0}, \ldots, c_{n,l}$ such that $R_q(c_{n,i})$ holds if and only if the $i$th state of the computation is $q$. Then the automaton $\mathcal{A}$ accepts $w$ if and only if:

- $R_{q_0}(c_{n,0})$,
- $R_q(c_{n,i})$ implies $R_{\delta(q,w[i])}(c_{n,i+1})$ and
- $R_q(c_{n,l})$ for some $q \in F$.

The precise values for the $c_{n,i}$'s are now given.

For $i \in [l]$, let $w_{\geq i}$ be the suffix of $w$ of length $l - i$, that is, $w$ without its $i$ first letters. It is the part of the word that remains to be read after the $i$th step of the computation of $\mathcal{A}$. Then let $c_{n,i}$ be $\overline{w_{\geq i}}^{\mathbb{N}}$. It implies that $c_{n,i+1} = \frac{c_{n,i} - w[i]}{b}$ and $w[i] = c_{n,i} \mod b$. hence, by an easy induction over $i$, $c_{n,i} = \lfloor \frac{n}{b^i} \rfloor$, and in particular $c_{n,0} = n$ and $c_{n,l} = 0$.

Similarly, let $q_i$ be the $i$th step of the computation of the automaton. That is, for all $i < l$:

$$q_{i+1} = \delta(q_i, w[i]) = \delta(q_i, c_{n,i} \mod b).$$

For all $q \in Q$, let $R_q^{\mathcal{S}}$ be $\{c_{n,i} \mid q_i = q\}$. The sets $R_q^{\mathcal{S}}$ can be defined by induction over $i$ as follows: $n \in R_{q_0}^{\mathcal{S}}$ and, for all $c_{n,i} \in R_q^{\mathcal{S}}$, if $c_{n,i} \equiv a \mod b$ then $c_{n,i+1} = \frac{n-a}{b} \in R_{\delta(q,a)}$.

Formally, we introduce the formula $\psi_\delta$ which states that two consecutive states encode a correct step of the computation. If $x$ is interpreted by $c_{n,i+1}$ and $q$ by $q_i$ then the formula holds only if $a$ is interpreted by $w[i]$ and $b \times x + a$ by $c_{n,i}$. Hence $\psi_\delta$ asserts that $R_{q_i}(c_{n,i})$ implies that $R_{\delta(q_i,w[i])}(c_{n,i+1})$:

$$\psi_\delta = (\forall x. x \not\doteq 0) \bigwedge_{q \in Q} \bigwedge_{a=0}^{b-1} R_q(b \times x + a) \Longrightarrow R_{\delta(q,a)}(x).$$

We want to ensure that $c_{n,0}$ corresponds to an initial state and $c_{n,l}$ corresponds to an accepting state. This can be expressed by the formula $\psi_{0,F}$:

$$\psi_{0,F} = R_{q_0}(\text{last}) \wedge \bigvee_{q \in F} R_q(0).$$

Finally, let $\psi_Q$ be the formula that states that there is at most one state by letter.

$$\psi_Q = \forall x. \bigwedge_{q \neq q' \in Q} \neg \left[R_q(x) \wedge R_{q'}(x)\right].$$

The three formulas $\psi_\delta$, $\psi_{0,F}$ and $\psi_Q$ belong to $\Pi_2[+1, \times b, (R_q)_{q \in Q}]$. Hence their conjunction $\psi = \psi_\delta \wedge \psi_{0,F} \wedge \psi_Q$ also belongs to $\Pi_2[+1, \times b, (R_q)_{q \in Q}]$. And by construction $\mathsf{SP}(\psi) = \overline{\mathcal{A}}^{\mathbb{N}}$. □

Example 31 is now resumed.

**Example 32.** For the set $\{2^i \mid i \in \mathbb{N}\}$, the formula $\psi$ is:

$$\begin{array}{l}
\exists R_{q_0}, R_{q_1}, R_{q_2}. \wedge R_{q_0}(\text{last}) \wedge R_{q_1}(0) \wedge (\forall x. x \not\doteq 0) \{ \\
R_{q_0}(2 \times x) \Longrightarrow R_{q_0}(x) \wedge R_{q_0}(2 \times x + 1) \Longrightarrow R_{q_1}(x) \wedge \\
R_{q_1}(2 \times x) \Longrightarrow R_{q_2}(x) \wedge R_{q_1}(2 \times x + 1) \Longrightarrow R_{q_2}(x) \wedge \\
R_{q_2}(2 \times x) \Longrightarrow R_{q_2}(x) \wedge R_{q_2}(2 \times x + 1) \Longrightarrow R_{q_2}(x)\} \wedge \\
\forall x. \{\neg \left[R_{q_0}(x) \wedge R_{q_1}(x)\right] \wedge \neg \left[R_{q_0}(x) \wedge R_{q_2}(x)\right] \wedge \neg \left[R_{q_2}(x) \wedge R_{q_1}(x)\right]\}.
\end{array}$$

Theorem 29 is now proven.

*Proof.* Let $S \subseteq \mathbb{N}^{>0}$ be a $b$-recognizable set. By Lemma 30, it suffices to prove that $S-1$ is accepted by an automaton in base $b$. Since $S \in b-\text{REC}$, by Theorem 24, there exists a $\mathsf{FO}[+, V_b]$-formula $\sigma(x)$ that defines $S$ (in $\mathbb{N}$). Then $\sigma(x+1)$ defines $S-1$. Hence $S-1$ is also $b$-recognizable. By definition of $b$-recognizable sets, there exists a deterministic automaton $\mathcal{A} = (Q, [b-1], \delta, q_0, F)$ that accepts $\underline{S-1}_{\mathbb{N}}$. □

Example 33 is now resumed.

**Example 33.** Let $S = \{2^i + 1 \mid i \in \mathbb{N}\}$, then $S - 1 = \{2^i \mid i \in \mathbb{N}\}$ and $S$ is the spectra of Formula (1) of Example 31.

Theorem 29 admits the following corollary.

**Corollary 34.** *The following inclusion holds:*

$$b - \textit{REC} \subseteq \mathsf{SP}\left(\forall \mathsf{MSO}\Sigma_2^2[+1, \times b]\right).$$

*Proof.* Let $S \subseteq \mathbb{N}^{>0}$ be a $b$-recognizable set. The set $\mathbb{N}^{>0} \setminus S$ is also $\mathsf{FO}[+, V_b]$-definable. Hence $\mathbb{N}^{>0} \setminus S$ is $b$-recognizable. By Theorem 29, there exists $\psi \in \exists\mathsf{MSO}\Pi_2^2[+1, \times b]$ such that $\psi^{\mathbb{N}} = \mathbb{N}^{>0} \setminus S$. Then $\mathsf{SP}(\neg\psi) = S$ and $\neg\psi \in \forall\mathsf{MSO}\Sigma_2^2[+1, \times b]$. □

It should be noted that, in general, $\mathsf{SP}(\neg\varphi)$ is not equal to $\mathbb{N}^{>0} \setminus \mathsf{SP}(\varphi)$. In our context, the equality holds because the vocabulary contains only interpreted relations.

As a corollary of Theorem 29, the same result is proven over words.

**Corollary 35.** *The class $b-$REC is a subset of the set of $\Pi_2^2[+1, \times b]$-spectra over words over some alphabet.*

That is, $b-$REC is a subset of the $\mathsf{FO}^2[+1, \times b]$-spectra.

*Proof.* Let $S \subseteq \mathbb{N}^{>0}$ be a $b$-recognizable set. By Theorem 29 there exists a $\exists\mathsf{MSO}\Pi_2^2[+1, \times b]$-formula $\varphi$ such that $\mathsf{SP}(\varphi) = S$. By construction $\varphi$ is of the form $\exists R_0, \ldots, R_{d-1}.\psi$ with $\psi \in \Pi_2^2[+1, \times b, R_0, \ldots, R_{d-1}]$. By Lemma 16, $\mathsf{SP}(\psi) = S$.

Let us define the alphabet $\mathcal{A}$ as the set of subsets of $[d-1]$. Let $\psi'$ be obtained from $\psi$ by replacing each atomic formula of the form $R_i(t)$ by $\bigvee_{\substack{I \subseteq [d-1] \\ i \in I}} P_I(t)$. Let us prove that $\mathsf{SP}(\psi') = \mathsf{SP}(\psi)$, that is, $\mathsf{SP}(\psi') = S$.

- Assume first that $n \in \mathsf{SP}(\varphi)$, then there exists a $\{+1, \times b, (R_i)_{i\in[d-1]}\}$-model $\mathcal{S}$ of $\psi$. The word $w_{\mathcal{S}}$ of length $n$ with $w_{\mathcal{S}}[j] = \{i \in [d-1] \mid S \models R_i(j)\}$ is clearly a model of $\psi'$. Hence $n \in \mathsf{SP}(\psi')$.
- Conversely, let $w \in \mathcal{A}^0$ be a model of length $n$ of $\psi'$. Let $\mathcal{S}$ be the $\{[n-1], +1, \times b, (R_i)_{i\in[d-1]}\}$-structure such that $R_i^{\mathcal{S}} = \{j \in [n-1] \mid i \in w[j]\}$. Then $\mathcal{S}$ is a model of $\psi$, hence $n \in \mathsf{SP}(\varphi)$.

□

## 3.2 From spectra of non-deterministic 2-counter automata to $\exists\mathsf{MSO}^2[+1, \times b]$-spectra

This section considers spectra of non-deterministic 2-counter automata, that is, sets of possible halting times of such automata. Those spectra are encoded as $\exists\mathsf{MSO}^2[+1, \times b]$-spectra. More generally, they are encoded as $\exists\mathsf{MSO}^2[+1, g]$ for a class of unary functions $g$ which contains $\times b$ for $b \geq 2$.

In this section, we encode a non-deterministic 2-counter automaton as a $\{+1, g\}$-structure for $g$ belonging to a large class $\mathbb{I}$ of functions. This implies that the satisfiability problem of this logic is undecidable for any function $g \in \mathbb{I}$.

The following definition introduces the class $\mathbb{I}$ of functions $g$ and some related notations.

**Definition 36.** Let $\mathbb{I}$ be the set of increasing functions $g : \mathbb{N} \to \mathbb{N}$ such that $g(1) > 1$ and such that the image of $g$ is not co-finite - that is, for each $n \in \mathbb{N}$, there is $s > n$ such that $g(s+1) > g(s) + 1$.

Let $g \in \mathbb{I}$. Let $\cong^g$ be the equivalence relation such that $n_0 \cong^g n_1$ if and only if ($g^i(n_0) = n_1$ or $g^i(n_1) = n_0$) for some $i \in \mathbb{N}$, where $g^i$ denotes the function $g$ applied $i$ times.

Then let $E^g(n)$ be the equivalence class of $n$ and $e^g(n)$ be the least integer of $E^g(n)$.

Let $(s_n^g)_{n\in\mathbb{N}}$ and $(p_n^g)_{n\in\mathbb{N}}$ be two sequences defined recursively by:

- $s_0^g = 1$.
- for all $i \in \mathbb{N}$ such that $s_i^g$ is defined:
  - $p_i^g$ is the least integer, greater than $s_i^g$, which has no antecedent by $g$. Formally:
$$p_i^g = \min\{n \in \mathbb{N} \mid n > s_i^g, \neg\exists m.g(m) = n\}.$$
  - $s_{i+1}^g$ is the least integer of $E^g(1)$ greatest than $p_i^g$. Formally
$$s_{i+1}^g = \min\{n \in \mathbb{N} \mid n > p_i^g, e^g(n) = 1\}.$$

We give some examples of such functions $g$'s and of notations introduced in Definition 36.

**Example 37.**

- Let $g$ denote the function $\times b$ for $b \geq 3$. The function $g$ belongs to $\mathbb{I}$. We have $e^g(n) = \frac{n}{V_b(n)}$, and the equivalence classes have the form $\{p \times b^n \mid n \in \mathbb{N}\}$ where $b$ does not divide $p$. Moreover $s_n^g = b^n$ and finally $p_n^g = b^n + 1$.

  The case $b = 2$ is almost identical, except that $s_n^g = 2^{n+1}$ for $n \in \mathbb{N}^{>0}$ and $p_n^g = 2^{n+1} + 1$ for $n \in \mathbb{N}$. The following array gives the first values for $g(n) = 2 \times n$.

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $g(n)$ | 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 | 18 | 20 | 22 | 24 | 26 | 28 | 30 | 32 | 34 | 36 | 38 | 40 |
| $e^g(n)$ | 1 | 1 | 3 | 1 | 5 | 3 | 7 | 1 | 9 | 5 | 11 | 3 | 13 | 7 | 15 | 1 | 17 | 9 | 19 | 5 |
| | $s_0^g$ | | $p_0^g$ | $s_1^g$ | $p_1^g$ | | | $s_2^g$ | $p_2^g$ | | | | | | | $s_3^g$ | $p_3^g$ | | | |

- Now let $g(n) = n + \lfloor \sqrt{n} \rfloor$. The first values for this function are:

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $g(n)$ | 2 | 3 | 4 | 6 | 7 | 8 | 9 | 10 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 20 | 21 | 22 | 23 | 24 |
| $e^g(n)$ | 1 | 1 | 1 | 1 | 5 | 1 | 5 | 1 | 5 | 1 | 11 | 5 | 1 | 11 | 5 | 1 | 11 | 5 | 19 | 1 |
| | $s_0^g$ | | | | $p_0^g$ | $s_1^g$ | | | | | $p_1^g$ | | $s_2^g$ | | | | | | $p_2^g$ | $s_3^g$ |

  Then the $\cong^g$-classes are $\{0\}$, $\{1, 2, 3, 4, 6, 8, 10, 13, 16, 20, \dots\}$, $\{5, 7, 9, 12, 15, 18, \dots\}$, $\{11, 14, 17, \dots\}$, $\{19, \dots\}, \dots$. The sequence $(p_n^g)$ begins by: $(5, 11, 19, \dots)$ (which is actually equal to $\{n^2 + n - 1 \mid n \in \mathbb{N}\}$), and the sequence $(s_n^g)$ begins by $(1, 6, 10, 20, \dots)$.

Note that the proof of the following theorem is based on the same technique as the proof of Theorem 4.4 of [16], but is more involved.

We now give a relation between spectra over non-deterministic 2-counter automata and over formulas.

**Theorem 38.** *For each non-deterministic 2-counter automaton $\mathcal{A}$, there exists a formula $\varphi_{\mathcal{A}} \in \exists \mathsf{MSO}\Pi_2^2[+1, g]$ such that if $g \in \mathbb{I}$ then $\mathsf{SP}(\varphi_{\mathcal{A}}) = \{s_t^g + 1 \mid t \in \mathsf{SP}(\mathcal{A})\}$.*

*Proof.* We encode a computation $(\kappa, c_0, c_1)$ of a non-deterministic 2-counter automaton $\mathcal{A}$ with a formula $\varphi_{\mathcal{A}}$ in $\exists \mathsf{MSO}\Pi_2^2[+1, g]$ such that $[s_t^g] \models \varphi_{\mathcal{A}}$ if and only if $\mathcal{A}$ admits a computation that halts in $t$ steps.

The formula $\varphi_{\mathcal{A}}$ will be the conjunction of several formulas which are defined below.

Let us assume that the computation halts after $t$ steps, then we divide our finite structure of universe $[s_t^g]$ in $t$ segments. The $i$th segment $S_i = [s_i^g, s_{i+1}^g - 1]$ encodes the $i$th step of the computation.

More precisely, we use monadic second order variables:

- The variable $S$ stands for the set of $s_i^g$'s.
- The variable $C_j$ encodes the $j$th counter. More precisely, in a segment $S_i$, a $\cong^g$-class $E$ is either entirely contained in $C_j$ or disjoint from $C_j$, namely $E \cap S_i \cap C_j \in \{\emptyset, E \cap S_i\}$. And the number of $\cong^g$-classes that intersect $S_i \cap C_j$ is equal to $c_j[i]$. Formally, $\#\{E^g(n) \mid n \in S_i \cap C_j\} = c_j[i]$.
- The variables $(Q_q)_{q \in [\#\mathcal{A}-1]}$ stand for the instruction number, that is: $Q_q$ is true on the (whole) segment $S_i$ if and only if $\kappa[i] = q$. In particular, there exists one and only one state by segment, the first state is 0 and the last is $\#\mathcal{A} - 1$. This can be expressed by the formula:

$$\varphi_Q = Q_0(1) \wedge Q_{\#\mathcal{A}-1}(\text{last}) \wedge \forall x \neq 0. \bigvee_{q=0}^{\#\mathcal{A}-1} \left\{ Q_q(x) \wedge \bigwedge_{q,q' \in Q, |q \neq q'} \neg Q_{q'}(x) \wedge [S(x+1) \vee Q_q(x+1)] \right\}.$$

The set $S$, the counters $C_j$ and the states $Q_q$ cannot be directly defined. Hence we introduce auxiliary second order variables. We then give the intended interpretation of those variables and their formal definitions. Finally, we define $S$, $C_0$, $C_1$ and the $Q_q$'s using those auxiliary variables.

**Defining $S$** The set $S$ cannot be defined directly, since its definition use the order relation, $E^g(1)$ and the sequence $(p_n^g)$. Hence we use the following second order variables:

- The variable $E$ stands for $E^g(1)$; formally $n$ belongs to $E$ if $n = 1$ or if there exists $n' \in \mathbb{N}$ such that $n' \in E$ and $g(n') = n$. Formally it is characterized by:

$$\varphi_E = \forall x.E(x) \Longleftrightarrow \{x \doteq 1 \vee [x \neq\!\!\!\doteq 1 \wedge (\exists y.g(y) \doteq x \wedge E(y))]\}.$$

- The variable $P$ stands for the union of the $[s_i^g, p_i^g - 1]$. Formally it is characterized by:

$$\varphi_P = \forall x.P(x) \Longleftrightarrow \{E(x) \vee [x \neq\!\!\!\doteq 0 \wedge P(x-1) \wedge \exists y.g(y) \doteq x]\}.$$

- Finally $S$ can be characterized by the set of letters of $E$ whose predecessor does not belong to $P$:

$$\varphi_S = \forall x.\left\{S(x) \Longleftrightarrow [E(x) \wedge \neg P(x-1)]\right\}.$$

Since we require that the cardinality of the structure has the form $s_t^g + 1$, we impose that $\mathrm{last} \in S$. Let

$$\varphi_{\mathrm{last}} = S(\mathrm{last}).$$

**Counters** The main issue with this encoding concerning the counters is that, in order to ensure that two successive segments encode two successive configurations, we need to compare the cardinality of two sets. And comparing cardinality does not seem possible in our logic. To overcome this, we use the property of the function $g$. If the $i$th step is a jump or a jz instruction, then $c_j[i] = c_j[i+1]$, and in this case we can choose $C_j$ in $S_{i+1}$ to be $\{n \in S_{i+1} \mid \exists m \in C_j.g(m) = n\}$. If the $i$th step is $incr(j)$ then it suffices to add a single position to $C_j$ in $S_{i+1}$. Such a position exists by construction, and equals $p_{i+1}^g$. It is a single position since $p_{i+1}^g$ has an antecedent and $g(p_{i+1}^g)$ belongs to $S_{i+2}$. Finally, if the $i$th step is $decr(j)$ then it suffices to remove the least integer of the image of $C_j$ which belongs to $S_i$.

More formally:

- If the $i$th step is $incr(j)$, that is $\kappa[i] = q$ with $\mathcal{A}_q = incr(j)$ and $Q_q(s_i^g)$, then:
  - $C_j(p_{i+1}^g)$,
  - $C_{1-j}(p_{i+1}^g)$ does not hold, and
  - for every $n \in \left[s_{i+1}^g, s_{i+2}^g - 1\right] \setminus \left\{p_{i+1}^g\right\}$, for $k \in \{0,1\}$, the property $C_k(n)$ holds if and only if there exists $n'$ with $g(n') = n$ and $C_j(n')$ holds.

  We introduce a new second order variable $R_j^+$ stating that the $j$th counter should be incremented. It is true on $[s_{i+1}^g, p_{i+1}^g - 1]$:

$$\varphi_{R_j^+,q} = \forall x.R_j^+(x) \Longleftrightarrow \left\{[S(x) \wedge Q_q(x-1)] \vee \left[\exists y.g(y) = x \wedge R_j^+(x-1)\right]\right\}$$

  It should be noted that the formula is correct because $p_{i+1}^g > s_{i+1}^g$.

- If the $i$th step is $decr(j)$, let $r$ be the least integer of $S_i$ such that $C_j(r)$ holds if it exists, and be $s_{i+1}^g$ otherwise. Then:
  - for every $n \in \left[s_{i+1}^g, s_{i+2}^g - 1\right]$, $C_j(n)$ holds if and only if there exists $n'$ with $g(n') = n$ such that $C_j(n')$ holds and $n \neq r$.
  - for every $n \in \left[s_{i+1}^g, s_{i+2}^g - 1\right]$, $C_{1-j}(n)$ holds if and only if there exists $n'$ with $g(n') = n$ such that $C_{1-j}(n')$ holds.

We must define $r$: for this, we introduce a new second order variable $R_j^-$ that holds only for integers of the segment $\left[s_{i+1}^g, r-1\right]$:

$$\varphi_{R_j^-,q} = \forall x. R_j^-(x) \Longleftrightarrow \left\{ [S(x) \wedge Q_q(x-1)] \vee \left[\neg S(x) \wedge R_j^-(x-1) \wedge \neg \left(\exists y. g(y) \doteq x \wedge C_j(y)\right)\right]\right\}.$$

- In every other case, $n \in \left[s_{i+1}^g, s_{i+2}^g - 1\right]$, $C_j\,(n)$ holds if and only if there exists $n'$ with $g\,(n') = n$ and $C_j\,(n')$ holds.

The $C_j$ are finally defined by

$$\varphi_{C_j} = \forall x. C_j(x) \Longleftrightarrow \left\{ \left[R_j^+(x-1) \wedge \neg \exists y. g(y) = x\right] \vee \left[\neg R_j^-(x-1) \wedge \exists y. g(y) \doteq x \wedge C_j(y)\right]\right\}$$

Let us prove that $C_k(n)$ does not hold for any $n \in E^g\,(1)$. Let (*) denote this statement. For the sake of contradiction, let $n \cong^g 1$ be the least integer such that $C_k(n)$ holds. If $n = 1$, the property (*) does not hold since $c_k[0] = 0$. If $n > 1$ then either $n = p_i^g$ or there exists $n'$ such that $g(n') = n$ and $C_k(n')$. The first case is impossible since $p_i^g \ncong^g 1$ by definition on $p_i^g$, and the second case contradicts the minimality of $n$.

**States** Finally, the formula states that the different states appear in the correct order. We have already stated in $\varphi_Q$ that the initial state is 0 and that the last state is $\#\mathcal{A} - 1$. It remains to encode that two successive segments encode a step of the computation.

Let $i \in [t-1]$ be a step, such that at the $i$th step of the computation the state of the automaton is $q$.

- If $\mathcal{A}_q = jmp\,(q_0, q_1)$, then either $Q_{q_0}\left(s_{i+1}^g\right)$ or $Q_{q_1}\left(s_{i+1}^g\right)$. This transition can be defined by:

$$\varphi_q = \forall x.\, [S(x) \wedge Q_q(x-1)] \Longrightarrow \{Q_{q_0}(x) \vee Q_{q_1}(x)\}.$$

- If $\mathcal{A}_q = jz\,(j, q')$, then if there exists a position $z_{i,j} \in \left[s_i^g, s_{i+1}^g - 1\right]$ such that $C_j\,(z_{i,j})$ then $Q_{q+1}(s_{i+1}^g)$, otherwise $Q_{q'}(s_{i+1}^g)$.

  This condition cannot be stated directly, as it uses the order relation. Hence we introduce an auxiliary variable $Z_j$ to determine whether the value of the $j$th counter equals 0 or not in the segment. More precisely, $Z_j$ is such that $Z_j \cap S_i = [s_i^g, z_{i,j} - 1]$ for the minimal value of $z_{i,j}$, if such a value exists, and $Z_j \cap S_i = S_i$ otherwise. Formally:

$$\varphi_{Z_j} = \forall x. Z_j(x) \Longleftrightarrow \{S(x) \vee [Z_j(x-1) \wedge \neg C_j(x)]\}.$$

  Then the transition can be defined by:

$$\varphi_q = \forall x.\, [S(x) \wedge Q_q(x-1)] \Longrightarrow \{[Z_j(x-1) \wedge Q_{q'}(x)] \vee [\neg Z_j(x-1) \wedge Q_{q+1}(x)]\}$$

- If $\mathcal{A}_q = incr(j)$ or $decr(j)$, that is $Q_q\,(s_i^g)$ with $\kappa\,(q)$ being $incr(j)$ or $decr(j)$, then $Q_{q+1}\left(s_{i+1}^g\right)$.

$$\varphi_q = \forall x.\, [S(x) \wedge Q_q(x-1)] \Longrightarrow Q_{q+1}(x)$$

Finally $\varphi_{\mathcal{A}}$ is defined by the $\exists \mathsf{MSO}\Pi_2\,[+1, g]$-formula:

$$\exists C_0, C_1, R_0^+, R_1^+, R_0^-, R_1^-, Z_0, Z_1, E, S, P, (Q_q)_{q\in[\#\mathcal{A}-1]}.\varphi_Q \wedge \varphi_{\text{last}} \wedge$$
$$\left(\bigwedge_{q\in[\#\mathcal{A}-1]} \varphi_q\right) \wedge \varphi_E \wedge \varphi_P \wedge \varphi_S \wedge \left[\varphi_{C_j} \wedge \varphi_{Z_j} \wedge \bigwedge_{j=0}^{1} \left(\bigwedge_{\mathcal{A}_{[q]}=incr(j)} \varphi_{R_j^+,q}\right) \wedge \left(\bigwedge_{\mathcal{A}_{[q]}=decr(j)} \varphi_{R_j^-,q}\right)\right].$$

An incrementation step is represented in Figure 2. □

Combining the above theorem with Lemma 14 yields the following corollary.

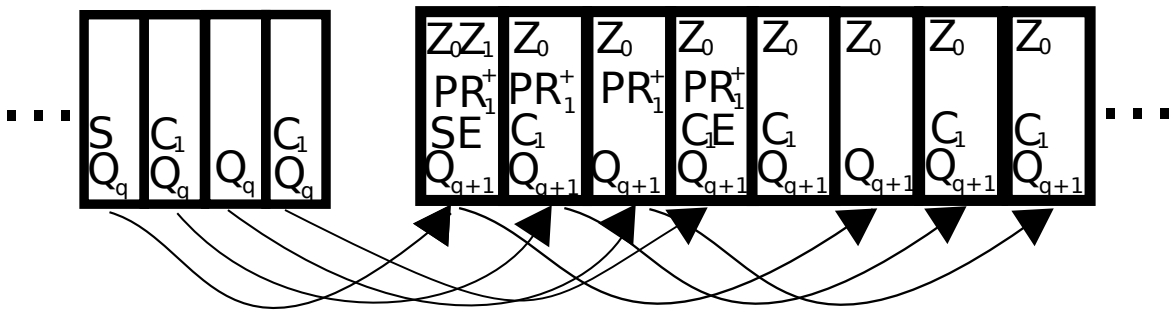


Figure 2: Example of a computation of $incr\,(1)$ in a non-deterministic 2-counter automaton

**Corollary 39.** *The class $b-$ REC is a strict subset of* $\mathsf{SP}\left(\exists\mathsf{MSO}\Pi_2^2\left[+1,\times b\right]\right)$.

*Proof.* The inclusion is proven in Theorem 29. Let us prove the strictness, that is $\mathsf{SP}\left(\exists\mathsf{MSO}\Pi_2^2\left[+1,\times b\right]\right)\not\subseteq b-\mathrm{REC}$. We use Theorem 38 with $g\left(n\right)=b\times n$, which has been studied in Example 37. If $\mathcal{A}$ is a non-deterministic 2-counter automaton whose spectrum is $S$, then the spectrum of $\varphi_{\mathcal{A}}$ is $\left\{b^n+1\mid n\in S\right\}$. Hence $\underline{\mathsf{SP}\left(\varphi_{\mathcal{A}}\right)}_{\mathbb{N}}$ is equal to $\left\{10^{n-1}1\mid n\in S\right\}$. Thus $\mathsf{SP}\left(\varphi_{\mathcal{A}}\right)$ is $b$-recognizable if and only if $S$ is ultimately periodic. Now, by Example 28, there exists a non-deterministic 2-counter automaton whose spectra is not ultimately periodic.

□

The next theorem states the undecidability result:

**Theorem 40.** *For $g\in\mathbb{I}$, the finite satisfiability of $\exists\mathsf{MSO}\Pi_2^2\left[+1,g\right]$, and the satisfiability of $\exists\mathsf{MSO}\Pi_2^2\left[+1,g\right]$ over $\mathbb{N}$, are undecidable.*

*Proof.* By Lemma 14 it suffices to prove the first claim. By Theorem 38, there exists a $\exists\mathsf{MSO}\Pi_2^2\left[+1,g\right]$-formula $\varphi_{\mathcal{A}}$ such that $\mathsf{SP}\left(\varphi_{\mathcal{A}}\right)\neq\emptyset$ if and only if $\mathcal{A}$ halts, which is undecidable [19]. □

This theorem admits the following corollary.

**Corollary 41.** *For every $g\in\mathbb{I}$, satisfiability of $\Pi_2^2\left[+1,g\right]$ over words is undecidable.*

The proof is exactly the same as the one of Corollary 35, hence it is not repeated.

# 4 First-order logic with successor and an uninterpreted function

The results of this section are similar to the results of the previous section, but concern the first-order logic with the successor function and an uninterpreted unary function symbol.

Section 4.1 is about spectra which are $*-$ REC. Section 4.2 is about spectra which encodes spectra of non-deterministic 2-counter automata. Finally, Section 4.3 deals with encodings of some extremely quickly increasing functions in $\mathsf{FO}[+1,f]$.

## 4.1 $b$-recognizable sets are $\mathsf{FO}[+1,f]$-spectra

We are going to prove a theorem similar to Theorem 29 for a logic with an uninterpreted function.

**Theorem 42.** *Every $*$-recognizable set is a $\Pi_2[+1,f]$-spectrum where $f$ is a uninterpreted unary function symbol.*

Let us emphasize the main differences between Theorem 29 and Theorem 42. Theorem 29 considers the vocabulary $\{+1,\times b\}$ while Theorem 42 considers $\{+1,f\}$. The former theorem considers monadic second-order logic with two first-order variables, while the latter considers first-order logic with any number of variables.

Uninterpreted function and two-variable logic are not considered together, because almost no restriction to the function can be expressed in two-variable logic. For example, if one wants to restrict $f$ to be the multiplication by 2, then it is naturally defined by $f(0) \doteq 0 \wedge \forall x.f(x) + 2 \doteq f(x+1)$, but $f(x) + 2 \doteq f(x+1)$ does not seem to be expressible in two-variable logic.

### 4.1.1 Examples

We first give two examples of 2-recognizable sets, and we show how to represent them as $\Pi_2[+1, f]$-spectra.

We first resume Example 33. This example shows how a formula can restrict the interpretation of $f$.

**Example 43.** Let $S = \{2^i + 1 \mid i \in \mathbb{N}\}$. We construct a $\{+1, f\}$-formula $\varphi$ whose spectrum is $S$. Let $\mathcal{S}$ be a $\{+1, f\}$-structure of universe $[c]$. If $\mathcal{S} \models \varphi$ then $c = 2^i$ for some $i \in \mathbb{N}$. Reciprocally, for all $i \in \mathbb{N}$, there exists a $\{+1, f\}$-structure $\mathcal{S}$ of universe $[2^i]$ such that $\mathcal{S} \models \varphi$. For the sake of simplicity, we assume that $c \geq 5$.

The formula $\varphi$ ensures that $f$ encodes the set of powers of 2, as well as multiplication by 2. The multiplication is encoded over even numbers, and the set of power of 2 is encoded over odd numbers.

Consider the function $g : [c] \to [c]$ defined as follows:

$$g(n) = \begin{cases} 2n & \text{if } n \text{ is even and } 2n \leq c \\ 1 & \text{if } n \text{ is even and } 2n > c \\ 3 & \text{if } n = 2^i - 1 \text{ for } i > 0 \\ 5 & \text{otherwise.} \end{cases}$$

Observe that $c$ is a power of 2, namely $c = 2^i$ for $i \in \mathbb{N}$, if and only if $g(\text{last} - 1) = 1$ (for $i > 0$) or $\text{last} = 1$ (for $i = 0$).

We now construct the formula $\varphi$ which ensures that the symbol $f$ is interpreted as $g$.

We first assert that for all $n > 0$, $f(n) \in \{3, 5\}$ if and only if $f(n-1) \notin \{3, 5\}$.

$$\forall x \neq 0.\, [f(x) \doteq 3 \vee f(x) \doteq 5] \Longleftrightarrow \{\neg\, [f(x-1) \doteq 3 \vee f(x-1) \doteq 5]\}.$$

We now assert that the interpretation of $f$ over even numbers is correct. This is true if the following conditions hold:

- $f(0) = 0$.
- For each $x$ such that $f(x) \notin \{1, 3, 5\}$ and $f(x) + 4 \leq c$, then $f(x+2) = f(x) + 4$. Indeed, in this case, $x$ is even, then $x+2$ is also even, and $f(x+2)$ should be $2 \times (x+2) = 2 \times x + 4 = f(x) + 4$. It can be asserted by:

$$\xi_{\leq \frac{c}{2}} = \forall x \{f(x) \neq 1 \wedge f(x) \neq 3 \wedge f(x) \neq 5\} \Longrightarrow \{ \quad [(\exists z.z \doteq f(x) + 4) \quad \Longrightarrow f(x+2) \doteq f(x) + 4] \\ \wedge [(\neg \exists z.z \doteq f(x) + 4) \Longrightarrow f(x+2) \doteq 1]\}.$$

- For each $x$ such that $f(x) = 1$ then $f(x+2) = 1$.

$$\xi_{> \frac{c}{2}} = \forall x.f(x) \doteq 1 \Longrightarrow f(x+2) \doteq 1$$

Let $\psi_{\text{even}}$ be the conjunction $f(0) \doteq 0 \wedge \xi_{\leq \frac{c}{2}} \wedge \xi_{> \frac{c}{2}}$.

Then we can use $2 \times x \doteq y$ as an abbreviation for

$$y \neq 1 \wedge y \neq 3 \wedge y \neq 5 \wedge \{f(x) \doteq y \vee f(x-1) \doteq y - 2\}.$$

The interpretation of $f$ over odd numbers is correct if: $f(x) = 3$ is equivalent to $\{x = 1$ or $f(\frac{x-1}{2}) = 3\}$. Indeed $2^1 - 1 = 1$ and $f(\frac{x-1}{2}) = 1$ implies that $\frac{x-1}{2} = 2^i - 1$ for some $i \in \mathbb{N}$, hence that $x = 2^{i+1} - 1$. It is asserted as follows:

$$\psi_{\text{odd}} = \forall x. f(x) \doteq 1 \Longleftrightarrow \{x \doteq 3 \vee \exists y. [2y + 1 \doteq x \wedge f(y) \doteq 3]\}.$$

Finally $S$ is the spectrum of the following formula:

$$\{f(\text{last} - 1) \doteq 1 \vee \text{last} \doteq 1\} \wedge \psi_{\text{even}} \wedge \psi_{\text{odd}}.$$

The following example is more involved. Let the basis $b$ be 2. We start from a 2-recognizable set $S$. Let $\mathcal{A}$ be an automaton which accepts $S$. We then give a formula $\varphi$ of spectrum $3S$. The $\Pi_2[+1, f]$-formula $\varphi$ asserts that the model of universe $[3c]$ encodes a computation of $\mathcal{A}$ over $\underline{c}_{\mathbb{N}}$ using the uninterpreted function.

**Example 44.** Let $S$ be the set of numbers $n$ such that the word $\underline{n}_{\mathbb{N}}$ has an even number of 1.

The set $S$ is accepted by the automaton $\mathcal{A} = (\mathbb{Z}/2\mathbb{Z}, \{0, 1\}, +, 0, \{0\})$.

We construct a $\Pi_2[+1, f]$-formula $\varphi$ whose spectra is $3S$. The formula $\varphi$ asserts that, over the different equivalence classes modulo 3, the formula $f$ encodes respectively:

- the equivalence classes modulo 3,
- the multiplication by 2 and
- the computation of the automaton $\mathcal{A}$.

Let $\mathcal{S}$ be a $\{+1, f\}$-structure of cardinality $c$ and let $w = \underline{c}_{\mathbb{N}}$. For technical reasons, we assume that $c \geq 3$. Let $Q_i$ be the $i$th state of the computation of $\mathcal{A}$ on $c$. That is $Q_i$ is the number of 1 in the first $i$ letters of $w$.

Let $g : [c] \to [c]$ be the function defined as follows.:

- $g(3n) = 0$,
- $g(3n + 1) = \begin{cases} Q_i + 1 & \text{if } n = \lfloor \frac{c}{2^i} \rfloor, i \in \mathbb{N}, \\ 3 & \text{otherwise,} \end{cases}$
- $g(3n + 2) = \begin{cases} 6 \times n & \text{if } n \leq \frac{c}{6} \\ 5 & \text{otherwise} \end{cases}$.

We now construct the formula $\varphi$ which ensures that the symbol $f$ is interpreted as $g$.

The interpretation of $f$ over multiples of 3 is correct if and only if the two following conditions hold: $f(0) = 0$ and for each $x \not\equiv 2$, $f(x) = 0$ if and only if $f(x + 3) = 0$.

It is formalized by $\psi_0$:

$$f(0) \doteq 0 \wedge (\forall x. x \not\equiv 2) [f(x) \doteq 0 \Longleftrightarrow f(x + 3) \doteq 0].$$

Using this characterization, $n \equiv i \mod 3$ can be expressed as $f(n - i) \doteq 0 \wedge n - i \not\equiv 2$.

The interpretation of $f$ over $3\mathbb{N} + 2 \cap [\frac{c}{6}]$ is correct if and only if the two following conditions hold: $f(2) = 0$ and for each $n \equiv 2 \mod 3$, $f(n + 3)$ is equal to $f(n) + 6$ if it exists and to 5 otherwise. It is formalized by $\xi_0$:

$$f(2) \doteq 0 \wedge \forall x \equiv 2 \mod 3. \{ \quad [(\exists y \doteq f(x) + 6) \Longrightarrow f(x + 3) \doteq f(x) + 6] \\ \wedge \quad [(\neg \exists y. y \doteq f(x) + 6) \Longrightarrow f(x + 3) \doteq 5]\}.$$

Then $2 \times x = y$ can be expressed as $\bigvee_{i=0}^{2} x \equiv i \mod 3 \wedge f(x - i + 2) = y - 2 \times i$.

The interpretation of $f$ over $3\mathbb{N} + 2 \setminus [\frac{c}{6}]$ is formalized by $\xi_1$:

$$\forall n. f(n) \doteq 5 \Longrightarrow f(n + 3) \doteq 5.$$

Then the interpretation of $f$ over $3\mathbb{N} + 2$ is formalized by $\psi_1 = \xi_0 \wedge \xi_1$.

The interpretation of $f$ over $3\mathbb{N}+1$ is correct if and only if the two following conditions hold:

- $f(\text{last}) \doteq 1$, indeed: $f(\text{last}) = f(c) = f(\lfloor \frac{c}{2^0} \rfloor) = Q_0 + 1 = 0 + 1 = 1$.
- the value of $f(3n+1)$ is 1, 2 or 3. It is formalized by $\chi_0$:

$$\forall x \equiv 1 \mod 3. \bigvee_{i=1}^{3} f(x) \doteq i$$

- if $f(3n+1) = 1$ and $n$ is even then $f(3\frac{n}{2}+1) = 1$. Indeed, it implies that $n = \lfloor \frac{c}{2^i} \rfloor$ and $Q_i = 0$ for some $i \in \mathbb{N}$. Since $n_i$ is even, the $(i+1)$th read letter is a 0, then $Q_{i+1} = 0$, hence one has $f(3\frac{n}{2}+1) = f(3\lfloor \frac{c}{2^{i+1}} \rfloor + 1) = 1$.

  More generally, if $f(3n+1)$ is 1 or 2, if $n$ is even then $f(3\frac{n}{2}+1) = f(3n+1)$ and if $n$ is odd then $f(3\frac{n-1}{2}+1) = 3 - f(3n+1)$.

  It is formalized by the formula $\chi_1$:

$$\forall x \equiv 0 \mod 3 \bigwedge_{i=1}^{2} \{f(x+1) \doteq i\} \Longrightarrow$$
$$\{\exists y. [2 \times y \doteq x \wedge f(y+1) \doteq i] \vee [2 \times y + 1 \doteq x \wedge f(y+1) \doteq 3 - i]\}$$

Then let $\psi_2$ be the formula $f(\text{last}) \doteq 1 \wedge \chi_0 \wedge \chi_1$.

The computation is accepting if the last state is 0, that is $f(1) = 1$.

Finally, $3S$ is the spectrum of the formula $\varphi$ defined as follows

$$\psi_0 \wedge \psi_1 \wedge \psi_2 \wedge f(1) \doteq 1.$$

We show in the next section that, using properties of $b$-recognizable sets, we can replace $3S$ by $S$ in the spectrum.

### 4.1.2 General case

We now explain how to encode a boolean combination of $b$-recognizable sets as a $\Pi_2[+1, f]$-spectrum.

We introduce a class of functions $\Xi_{m,d,e}$ from $\{+1, (\times b)_{2 \leq b \leq m}\}$-structures to $\{+1, f\}$-structures and a function $\Phi_{m,d,e}$ from $\exists \mathsf{MSO}[+1, (\times b)_{2 \leq b \leq m}]$-formulas to $\mathsf{FO}[+1, f]$-formulas. The function $\Xi_{m,d,e}$ multiplies the cardinality by $m$. Moreover those functions are chosen such that $\mathcal{S} \models \varphi$ if and only if $\Xi_{m,d,e}(\mathcal{S}) \models \Phi_{m,d,e}(\varphi)$. Hence, if $\varphi$ has a model, then $\Phi_{m,d,e}(\varphi)$ also has a model. Conversely, if $\Phi_{m,d,e}(\varphi)$ has a model and this model belongs to the image of $\Xi_{m,d,e}$ then $\varphi$ also has a model. This properties impose to choose $\Xi_{m,d,e}$ such that there exists a $\mathsf{FO}[+1, f]$-formula which characterizes its image.

The function $\Phi_{m,d,e}$ is defined by induction over the $\exists \mathsf{MSO}[+1, (\times b)_{b \in [2,m]}]$-formulas. Let $\varphi \in \exists \mathsf{MSO}[+1, \times b]$ be of the form $\exists R.\psi$, then $\psi \in \exists \mathsf{MSO}[+1, \times b, R]$, hence $\Phi_{m,d,e}$ and $\Xi_{m,d,e}$ must be defined on vocabularies which extends $\{+1, \times b\}$ with a finite number of first and second order variables.

We now begin the formal definition.

**Definition 45.** Let $m > 1$ and let $d, e \in \mathbb{N}$. Let :

$$\mathcal{V}_{m,d,e} = \left\{+1, (\times b)_{b \in [2,m-1]}, (R_i)_{i \in [d-1]}, (x_i)_{i \in [e-1]}\right\},$$

and let

$$\mathcal{V}'_e = \left\{+1, f, (x_i)_{i \in [e-1]}\right\}.$$

We must now define the function $\Xi_{m,d,e}$ from $\mathcal{V}_{m,d,e}$ to $\mathcal{V}'_e$.

Semantically, $\Xi_{m,d,e}$ multiplies every value by $m$, and encodes every relation, apart from $+1$, with the unary function $f$. The image by $f$ of each equivalence class modulo $m$ encodes either the set of second order variables, or a multiplication, or the modular classes.

**Definition 46.** The function $\Xi_{m,d,e}$ sends each $\mathcal{V}_{m,d,e}$-structure $\mathcal{S}$ of cardinality $\kappa$ to a $\mathcal{V}'_e$-structure $\Xi_{m,d,e}(\mathcal{S})$ of cardinality $m \times \kappa$, defined as follows:

- For $i \in [e-1]$, let $x_i^{\Xi_{m,d,e}(\mathcal{S})} = m \times x_i^{\mathcal{S}}$.
- The function $f^{\Xi_{m,d,e}(\mathcal{S})}$ is defined as follows:
  - $f^{\Xi_{m,d,e}(\mathcal{S})}(m \times n) = 0$ (to encode the congruence classes)
  - $f^{\Xi_{m,d,e}(\mathcal{S})}(m \times n + 1) = \left( \sum_{\substack{i \in [d-1] \\ R_i^{\mathcal{S}}(n)}} 2^{i+1} \right) + 1$ (to encode the value of all unary relations over $n$)
  - for every $b \in [2, m-1]$, $f^{\Xi_{m,d,e}(\mathcal{S})}(m \times n + b) = \begin{cases} m \times n \times b & \text{if } m \times n \times b \text{ belongs to } [n \times \kappa - 1] \\ 1 & \text{otherwise} \end{cases}$ (to encode the multiplication by $b$)

The value of $f^{\Xi_{m,d,e}(\mathcal{S})}(m \times n + 1)$ encodes the set of second order variables whose interpretation in $\mathcal{S}$ hold on $n$. This is the only value of $\Xi_{m,d,e}(\mathcal{S})$ which depends on $\mathcal{S}$.

It should be noted that the antecedent of 0 by $f^{\Xi_{m,d,e}(\mathcal{S})}$ is $(m \times [\kappa - 1]) \cup [2, m-1]$. Hence $x \equiv 0 \mod m$ if and only if $\Xi_{m,d,e}(\mathcal{S}) \models \text{Mul}_m(x)$ where:

$$\text{Mul}_m(x) = f(x) \doteq 0 \wedge \bigwedge_{i=2}^{m-1} x \not\doteq i.$$

In the following examples, we resume Example 33 and we show the exact result of the application of those functions, with $m = 4$. Note that Example 43 is not an application of the functions $\Phi_{m,d,e}$ and $\Xi_{m,d,e}$; some properties used in the general case are omitted in Example 43 to simplify the formula.

**Example 47.** Let $m = 4$, $d = 1$ and $e = 1$, then $\mathcal{V}_{4,1,1} = \{+1, \times 2, \times 3, R_0, x_0\}$ and $\mathcal{V}'_1 = \{+1, f, x_0\}$.

Let $\mathcal{S}$ be a $\mathcal{V}_{m,d,e}$-structure of cardinality $2^c+1$ for $c \in \mathbb{N}$ such that: $x_0^{\mathcal{S}} = \text{last}^{\mathcal{S}} = 2^c$ and $R_0^{\mathcal{S}} = \{2^i \mid i \in [c] \setminus \{0\}\}$.

Then $\Xi_{m,d,e}(\mathcal{S})$ is the structure of cardinality $4 \times \left(2^i + 1\right) = 2^{i+2} + 4$ such that:

- $x_0^{\Xi_{m,d,e}(\mathcal{S})} = 2^{c+2} = \left(2^{c+2} + 3\right) - 3 = \text{last}^{\Xi_{m,d,e}(\mathcal{S})} - 3$,
- The function $f^{\Xi_{m,d,e}(\mathcal{S})}$ is defined as:
  - $f^{\Xi_{m,d,e}(\mathcal{S})}(4 \times n) = 0$ for $n \in [2^c]$,
  - $f^{\Xi_{m,d,e}(\mathcal{S})}(4 \times n + 1) = \begin{cases} 3 & \text{if } n \text{ is of the form } 2^j \\ 1 & \text{otherwise} \end{cases}$
  - $f^{\Xi_{m,d,e}(\mathcal{S})}(4 \times n + 2) = 8 \times n$ if $n \in \left[2^{c-3}\right]$, and is undefined otherwise,
  - $f^{\Xi_{m,d,e}(\mathcal{S})}(4 \times n + 3) = 12 \times n$ if $n \in \left[\lfloor \frac{2^{c-2}}{3} \rfloor\right]$ and is undefined otherwise,

Finally, we introduce the function $\Phi_{m,d,e}$ from $\mathsf{FO}[\mathcal{V}_{m,d,e}]$ to $\mathsf{FO}[\mathcal{V}'_e]$. Intuitively, this function translates the formulas such that the properties introduced in Definition 46 are satisfied.

**Definition 48** ($\Phi_{m,d,e}$)**.** The function $\Phi_{m,d,e}$ is defined recursively as follows:

| $\varphi$ | $\Phi_{m,d,e}(\varphi)$ | $\varphi$ | $\Phi_{m,d,e}(\varphi)$ | $\varphi$ | $\Phi_{m,d,e}(\varphi)$ | $\varphi$ | $\Phi_{m,d,e}(\varphi)$ |
|---|---|---|---|---|---|---|---|
| $\chi \vee \psi$ | $\Phi_{m,d,e}(\chi) \vee \Phi_{m,d,e}(\psi)$ | $\neg\psi$ | $\neg\Phi_{m,d,e}(\psi)$ | $x+1 \doteq y$ | $x+m \doteq y$ | $b \times x \doteq y$ | $f(x+b) \doteq y$ |
| $\exists x.\psi$ | $\exists x.\text{Mul}_m(x) \wedge \Phi_{m,d,e+1}(\psi)$ | $x \doteq y$ | $x \doteq y$ | $R_i(x)$ | $\bigvee_{\substack{I \subseteq [d] \\ i \in I}} \left[ f(x+1) \doteq \sum_{j \in I} 2^{j+1} + 1 \right]$ | | |

We resume Example 47 to give an example of the application of $\Phi_{m,d,e}$ to the formula of Example 33.

**Example 49.** The formula $\varphi$ of Example 33 and the formula $\Phi_{m,d,e}(\varphi)$ are:

$$\begin{aligned}\varphi =& \exists R_{q_0}.R_{q_0}(\text{last}) \quad \wedge \neg R_{q_0}(0) \quad \wedge \forall x. \quad \{x \neq 0 \wedge \quad R_{q_0}(2\times x) \quad\} \wedge R_{q_0}(n') \quad .\\ \Phi_{m,d,e}(\varphi) =& \quad f(\text{last}-2) \doteq 3 \wedge \neg f(1) \doteq 3 \wedge \forall x.\text{Mul}_4(x)\{x \neq 0 \wedge f(f(x+2)+1) \doteq 3\} \wedge f(n'+1) \doteq 3 .\end{aligned}$$

We now prove a first lemma which states that the functions $\Xi_{m,d,e}$ and $\Phi_{m,d,e}$ preserve satisfiability.

**Lemma 50.** *Let $m, d, e \in \mathbb{N}$. Let $\varphi \in \mathsf{FO}[\mathcal{V}_{m,d,e}]$ and $\mathcal{S}$ be a $\mathcal{V}_{m,d,e}$-structure. Then $\mathcal{S} \models \varphi$ if and only if $\Xi_{m,d,e}(\mathcal{S}) \models \Phi_{m,d,e}(\varphi)$.*

*Proof.* By induction over $\varphi$:

- If $\varphi$ is $x_i \doteq x_j$, then the following statements are equivalent:

$$\begin{aligned}\mathcal{S} \models x_i \doteq x_j &\iff & x_i^{\mathcal{S}} &= x_j^{\mathcal{S}}\\ &\iff & m \times x_i^{\mathcal{S}} &= m \times x_j^{\mathcal{S}}\\ &\iff & x_i^{\Xi_{m,d,e}(\mathcal{S})} &= x_j^{\Xi_{m,d,e}(\mathcal{S})}\\ &\iff \Xi_{m,d,e}(\mathcal{S}) \models & x_i &\doteq x_j\\ &\iff \Xi_{m,d,e}(\mathcal{S}) \models \Phi_{m,d,e}( & x_i &\doteq x_j \quad ).\end{aligned}$$

- If $\varphi$ is $x_i + 1 \doteq x_j$, then the following statements are equivalent:

$$\begin{aligned}\mathcal{S} \models x_i + 1 \doteq x_j &\iff & x_i^{\mathcal{S}} + 1 &= x_j^{\mathcal{S}}\\ &\iff & m \times x_i^{\mathcal{S}} + m &= m \times x_j^{\mathcal{S}}\\ &\iff & x_i^{\Xi_{m,d,e}(\mathcal{S})} + m &= x_j^{\Xi_{m,d,e}(\mathcal{S})}\\ &\iff \Xi_{m,d,e}(\mathcal{S}) \models & x_i + m &\doteq x_j\\ &\iff \Xi_{m,d,e}(\mathcal{S}) \models \Phi_{m,d,e}( & x_i + 1 &\doteq x_j \quad ).\end{aligned}$$

- If $\varphi$ is $x_i \times b \doteq x_j$, then two cases must be considered, depending on the value of $x_i^{\mathcal{S}}$. Assume that $x_i^{\mathcal{S}} \geq \frac{\#\mathcal{S}}{b}$. Then $x_i^{\mathcal{S}} \times b \geq \#\mathcal{S}$, hence $x_i^{\mathcal{S}} \times b$ does not belong to the universe, which implies that $f\left(m \times x_i^{\mathcal{S}} + b\right) = 1$, which is not a multiple of $m$. Hence $\mathcal{S} \not\models f(x_i + b) \doteq x_j$. Hence $\mathcal{S} \models \varphi$ is trivially equivalent to $\Xi_{m,d,e}(\mathcal{S}) \models \Phi_{m,d,e}(x_i \times b \doteq x_j)$ since both statements are false.

  Now, assume that $x_i^{\mathcal{S}} < \frac{\#\mathcal{S}}{b}$, then the following statements are equivalent:

$$\begin{aligned}\mathcal{S} \models b \times x_i \doteq x_j &\iff & b \times x_i^{\mathcal{S}} &= x_j^{\mathcal{S}}\\ &\iff & b \times m \times x_i^{\mathcal{S}} &= m \times x_j^{\mathcal{S}}\\ &\iff & b \times x_i^{\Xi_{m,d,e}(\mathcal{S})} &= x_j^{\Xi_{m,d,e}(\mathcal{S})}\\ &\iff & f^{\Xi_{m,d,e}(\mathcal{S})}(x_i^{\Xi_{m,d,e}(\mathcal{S})} + b) &= x_j^{\Xi_{m,d,e}(\mathcal{S})}\\ &\iff \Xi_{m,d,e}(\mathcal{S}) \models & f(x_i + b) &\doteq x_j\\ &\iff \Xi_{m,d,e}(\mathcal{S}) \models \Phi_{m,d,e}( & b \times x_i &\doteq x_j \quad ).\end{aligned}$$

- If $\varphi$ is $R_i(x_j)$, then the following statements are equivalent:

$$\begin{aligned}\mathcal{S} \models R_i(x_j) &\iff \text{there exists } I \subseteq [d] \text{ containing } j \text{ such that } f^{\Xi_{m,d,e}(\mathcal{S})}\left(x_j^{\Xi_{m,d,e}(\mathcal{S})}\right) = \textstyle\sum_{k\in I} 2^{k+1} + 1\\ &\iff \Xi_{m,d,e}(\mathcal{S}) \models \bigvee_{\substack{i\in I\\ I\subseteq[d]}} f(x) = \textstyle\sum_{j\in I} 2^{j+1} + 1\\ &\iff \Xi_{m,d,e}(\mathcal{S}) \models \Phi_{m,d,e}(R_i(x_j)).\end{aligned}$$

- If $\varphi$ is $\exists x.\psi$, then it should be noted that $f^{\Xi_{m,d,e}(\mathcal{S})}(n) = 0$ if and only if ($n = p \times m + q$ with $q \in [m-1]$ and $q = 0$) or ($p = 0$ and $q > 1$). Then the following statements are equivalent:

$$\begin{array}{lll}
 & \mathcal{S} & \models \exists x.\psi \\
\Longleftrightarrow \text{there exists } c \in [s-1] \quad \text{such that} & \mathcal{S}\ [x/c] & \models \psi \\
\Longleftrightarrow \text{there exists } c \in [s-1] \quad \text{such that} & \Xi_{m,d,e}(\mathcal{S}\ [x/c]) & \models \Phi_{m,d,e}(\psi) \\
\Longleftrightarrow \text{there exists } c \in [s-1] \quad \text{such that} & \Xi_{m,d,e}(\mathcal{S})[x/m \times c] & \models \Phi_{m,d,e}(\psi) \\
\Longleftrightarrow \text{there exists } c' \in [m \times s - 1], \text{ a multiple of } m, \text{such that} & \Xi_{m,d,e}(\mathcal{S})[x/c'] & \models \Phi_{m,d,e}(\psi) \\
\Longleftrightarrow \text{there exists } c' \in [m \times s - 1] \quad \text{such that} & \Xi_{m,d,e}(\mathcal{S})[x/c'] & \models \Phi_{m,d,e}(\psi) \wedge \mathrm{Mul}_m(x) \\
\Longleftrightarrow & \Xi_{m,d,e}(\mathcal{S}) & \models \Phi_{m,d,e}(\varphi)\,.
\end{array}$$

- Finally if $\varphi$ is a negation or a disjunction, then the proof is a straightforward consequence of the induction hypothesis.

□

As explained in the introduction of this section, we need to characterize the image of $\Xi_{m,d,e}$. This is done in the following lemma.

**Lemma 51.** *Let $m, d, e \in \mathbb{N}$. There exists a $\Pi_2[+1, f]$-formula $\theta_{m,d,e}$ that is true over the $\mathcal{V}'_e$-structures belonging to the image of $\Xi_{m,d,e}$.*

*Proof.* We give a list of properties that characterize the image of $\Xi_{m,d,e}$, together with $\Pi_2[+1, f]$-formulas which express those properties.

- Let $\psi_0$ be the formula which states that the cardinality of the structure is at least $2^{d+1}$, using the formula introduced in Lemma 17:

$$\gamma_{\geq 2^{d+1}}.$$

- Let $\psi_1$ be the formula which states that all constants belong to $m\mathbb{N}$:

$$\bigwedge_{i=0}^{e-1} \mathrm{Mul}_m(c_i)$$

- Let $\psi_2$ be the formula which states that the cardinality of the structure is a multiple of $m$:

$$\mathrm{Mul}_m(\mathrm{last} - m + 1)$$

- For each $n \in m \times \mathbb{N}$:
  - Let $\xi_0$ be the formula which states that $f^{\Xi_{m,d,e}(\mathcal{S})}(n) = 0$:

$$f(n) \doteq 0.$$

  - Let $\xi_1$ be the formula which states that $f^{\Xi_{m,d,e}(\mathcal{S})}(n+1)$ is an odd integer between 1 and $2^{d+2} - 1$:

$$\bigvee_{i=0}^{2^d} f(n) \doteq 2i + 1.$$

  - For $q \in [2, m-1]$, let $\xi_{2,q}$ be the formula which states that $f^{\Xi_{m,d,e}(\mathcal{S})}(n+q)$ is $c \times n$ if $cn$ belongs to the universe, and equals 1 otherwise:

$$\begin{aligned} f(q) \doteq 0 \wedge \{[\exists y.y \doteq m \times q + f(n+q-m)] \Longrightarrow & f(n+q) \doteq m \times q + f(n+q-m)\} \\ \wedge \{\neg[\exists y.y \doteq m \times q + f(n+q-m)] \Longrightarrow & f(n+q) \doteq 1\}\,. \end{aligned}$$

    It should be noted that $m \times q$ and $q - m$ are constants, hence it is indeed a $\Pi_2[+1, f]$-formula.

– Let $\xi_3$ be the formula which states that if $f^{\Xi_{m,d,e}(\mathcal{S})}(n+q)=1$ for $q>1$, then $f^{\Xi_{m,d,e}(\mathcal{S})}(n+q+m)=1$:

$$\{f(n+q)\doteq 1\}\Longrightarrow\{f(n+m+q)\doteq 1\}.$$

Let $\psi_3$ be the formula

$$\forall n.\left(f(n)\doteq 0\wedge\bigwedge_{i=2}^{m-1}n\not\doteq i\right)\Longrightarrow\{\xi_0\wedge\xi_1\wedge\xi_{2,q}\wedge\xi_3\}$$

The image of $\Xi_{m,d,e}$ is then defined by the formula $\theta_{m,d,e}=\psi_0\wedge\psi_1\wedge\psi_2\wedge\psi_3$. □

Combining the formulas $\theta_{m,d,e}$ and $\Phi_{m,d,e}$ we can finally transform a $\exists\mathsf{MSO}[+1,(\times b)_{2\leq b<m}]$-spectrum $S$ into a $\mathsf{FO}[+1,f]$-spectrum $m\times S$.

**Lemma 52.** *Let $m\geq 2$, $d,e\in\mathbb{N}$, and $k<m$. Let $\varphi\in\mathsf{FO}[+1,(\times b)_{b\in[m-1]},(R_i)_{i\in[d-1]},(x_i)_{i\in[e-1]}]$. Let $\psi=\theta_{m,d,e}\wedge\Phi_{m,d,e}(\varphi)$. Then $\mathsf{SP}(\psi)=(m\times\mathsf{SP}(\varphi))\setminus[2^{d+1}-1]$.*

*Proof.* We prove the equivalence by proving the two inclusions.

Let us prove that $\mathsf{SP}(\psi)\subseteq(m\times\mathsf{SP}(\varphi))\setminus[2^{d+1}-1]$.

Let $n\in\mathsf{SP}(\psi)$. Then there exists a $\mathcal{V}'_e$-structure $\mathcal{S}'$ with cardinality $n$ such that $\mathcal{S}'\models\psi$. If follows from the definition of $\psi$ and Lemma 51 that the structure $\mathcal{S}'$ belongs to the image of $\Xi_{m,d,e}$. Hence $n$ is of the form $m\times n'$, with $n\geq 2^{d+1}$ and there exists a $\mathcal{V}_{m,d,e}$-structure $\mathcal{S}$ of cardinality $n'$ such that $\Xi_{m,d,e}(\mathcal{S})=\mathcal{S}'$. We have $\mathcal{S}'\models\Phi_{m,d,e}(\varphi)$, hence by Lemma 50, we have $\mathcal{S}\models\varphi$, which implies that $n'\in\mathsf{SP}(\varphi)$, hence $n\in m\times\mathsf{SP}(\varphi)$.

Let us prove that $\mathsf{SP}(\psi)\subseteq(m\times\mathsf{SP}(\varphi))\setminus[2^{d+1}-1]$.

Let $n\in(m\times\mathsf{SP}(\varphi))\setminus[2^{d+1}-1]$, then $n=m\times n'$ and $n\geq 2^{d+1}$ with $n'\in\mathsf{SP}(\varphi)$. We must prove that $n\in\mathsf{SP}(\psi)$. There exists a $\mathcal{V}_{m,d,e}$-model $\mathcal{S}$ of cardinality $n'$ such that $\mathcal{S}\models\varphi$. Hence by Lemma 50 $\Xi_{m,d,e}(\mathcal{S})\models\Phi_{m,d,e}(\varphi)$. Furthermore, by Lemma 51 we have $\Xi_{m,d,e}(\mathcal{S})\models\theta_{m,d,e}$. Hence $\Xi_{m,d,e}(\mathcal{S})$ is also a model of $\Phi_{m,d,e}(\varphi)\wedge\theta_{m,d,e}$, that is, a model of $\psi$. Moreover, by definition, the cardinality of $\Xi_{m,d,e}(\mathcal{S})$ is $n$. Hence $n\in\mathsf{SP}(\psi)$.

□

We can finally give the proof of Theorem 42.

Table 1 gives examples of values for the variables of the first part of the proof. It describes different ways to see the set $S=S_0\cap S_1$. The formulas $\zeta_i^k$ are not the ones given by Lemma 17, but some smaller equivalent ones. This simplifies the example without loss of generality. Indeed this proof is correct for any formula, and not only for the one generated by Lemma 17.

*Proof of Theorem 42.* Let $S$ be a $*$-recognizable set. Since $b$-recognizable sets are closed under complementation, it can be assumed that $S$ is a positive boolean combination of $b_{i,j}$-recognizable sets, that is, $S=\bigcup_i\bigcap_j S_{i,j}$ where every $S_{i,j}$ is $b_{i,j}$-recognizable for some $b_{i,j}\geq 2$. Let $m=\max(b_{i,j})+1$.

We create some intermediate sets and we prove that all such sets are $\mathsf{FO}[+,(V_b)_{b<m}]$-definable. We then reduce the problem to proving that those sets are $\Pi_2[+1,f]$-spectra.

By Lemma 18, spectra are closed by union, thus, in order to prove that $S$ is a $\Pi_2[+1,f]$-spectrum, it is sufficient to prove that every $S_i=\bigcap_j S_{i,j}$ is a $\Pi_2[+1,f]$-spectrum. Let $i$ be fixed in the remaining of this proof. Let $\chi_{i,j}(x)\in\mathsf{FO}[+,V_{b_{i,j}}]$ be a formula which defines $S_{i,j}$.

| $m$ | 4 |
|---|---|
| $b_0$ | 3 |
| $S_0$ | $\{3i+2\mid i\in\mathbb{N}\}$ |
| $\xi_0$ | $x\equiv 2 \mod 3\wedge x\geq 8$ |
| $S_0^k$ | $\{12\times n+8-3\times k\mid n\in\mathbb{N}\}$ for $k\in[4]$ |
| $T_0^k$ | $\{12\times n+8-4\times k\mid n\in\mathbb{N}\}$ for $k\in[4]$ |
| $U_0^k$ | $\{3\times n+2-k\mid n\in\mathbb{N}\}$ for $k\in[4]$ |
| $\zeta_0^k$ | $\exists(R_i)_{i\in[11]}.R_0(0)\wedge\forall x\not\doteq\text{last}.$ $\bigwedge_{i=0}^{11}[R_i(x)\Longleftrightarrow R_{i+1 \mod 12}(x+1)]$ $\wedge R_{16-3k \mod 12}(\text{last})$ |
| $b_1$ | 2 |
| $S_1$ | $\{2^i+1\mid i\in\mathbb{N}\}$ |
| $S_1^1$ | $\{2^i+1\mid i\in\mathbb{N}\}$ |
| $T_1^1$ | $\{2^i+1\mid i\in\mathbb{N}\}$ |
| $U_1^1$ | $\{2^i+1\mid i\in\mathbb{N}^{>0}\}$ |
| $\zeta_1^1$ | $\exists R_{q_0}.R_{q_0}(\text{last})\wedge\neg R_{q_0}(0)\wedge\forall n,n'.$ $[n\not\doteq 0\wedge R_{q_0}(n)\wedge 2\times n'\doteq n]\Longrightarrow R_{q_0}(n')$ |
| $S_1^k$ | $\emptyset$ for $k$ being 0, 2, 3 |
| $T_1^k$ | $\emptyset$ for $k$ being 0, 2, 3 |
| $U_1^k$ | $\emptyset$ for $k$ being 0, 2, 3 |
| $\zeta_1^k$ | false for $k$ being 0, 2, 3 |
| $S$ | $S_0\cap S_1=\{2^{2i}+1\mid i>1\}$ |
| $S^1$ | $\{2^{2i}+1\mid i>1\}$ |
| $T^1$ | $\{2^{2i}+1\mid i\in\mathbb{N}\}$ |
| $S^k$ | $\emptyset$ for $k$ being 0, 2, 3 |
| $T^k$ | $\emptyset$ for $k$ being 0, 2, 3 |

Table 1: Variables for a fixed set $S=S_0\cap S_1$ of Proof of Theorem 42

Let us prove that $S_i = \bigcap_j S_{i,j}$ is a $\Pi_2[+1, f]$-spectrum.

Since the function $\Xi_{m,d,e}$ multiplies the size of models by $m$, we partition $S$ into $m$ parts, one part for each congruence class. For $k \in [m-1]$, let $S^k_{i,j} = S_{i,j} \cap (m\mathbb{N} + k)$ and $S^k_i = S_i \cap (m\mathbb{N} + k)$. The set $S^k_{i,j}$ is $b$-recognizable as it is defined by $\chi^k_{i,j}(x) = \chi_{i,j}(x) \wedge x \equiv k \mod m$. The modular relations are $\mathsf{FO}[+]$-definable.

Then $S_i = \bigcup_{k\in[m-1]} S^k_i$. By Lemma 18 it is enough to prove that $S^k_i$ is a $\Pi_2[+1, f]$-spectrum. Let $k \in [m-1]$ be fixed for the remainder of the proof.

Let $T^k_{i,j} = S^k_{i,j} - k$, it is $b-$regular since it is defined by $\tau^k_{i,j}(x) = \chi^k_{i,j}(x + k)$.

Let $T^k_i = \cap_j T^k_{i,j}$. By Lemma 19, it is enough to prove that $T^k_i$ is a $\Pi_2[+1, f]$-spectrum. Let $U^k_{i,j} = \frac{T^k_{i,j}}{m}$. This set is also $b_{i,j}$-recognizable as it is defined by $\beta^k_{i,j}(x) = \tau^k_{i,j}(m \times x)$. By Theorem 29 there exist $d \in \mathbb{N}$ and an $\exists\mathsf{MSO}\Pi_2[+1, \times b_{i,j}]$-formula $\zeta^k_{i,j}$ of the form $\exists R_0, \ldots, R_{d-1}.\xi^k_{i,j}$ such that $\mathsf{SP}\left(\zeta^k_{i,j}\right) = U^k_{i,j}$. Without loss of generality, it can be assumed that no second order variable is used in two distinct formulas $\zeta^k_{i,j}$ and $\zeta^{k'}_{i',j'}$. Hence the formulas $\exists R_0, \ldots, R_{d-1}.\bigwedge_j \zeta^k_{i,j}$ and $\bigwedge_j \exists R_0, \ldots, R_{d-1}.\zeta^k_{i,j}$ are equivalent.

Let $\psi^k_i = \theta_{m,d,e} \wedge \Phi_{m,d,0}\left(\bigwedge_j \zeta^k_{i,j}\right)$, then by Lemma 52,

$$\begin{aligned}\mathsf{SP}\left(\psi^k_i\right) =& m \times \mathsf{SP}\left(\textstyle\bigwedge_j \zeta^k_{i,j}\right) \setminus [2^{d+1} - 1] = m \times [\textstyle\bigcap_j \mathsf{SP}\left(\zeta^k_{i,j}\right)] \setminus [2^{d+1} - 1] = \\ & m \times \left[\textstyle\bigcap_j U^k_{i,j}\right] \setminus [2^{d+1} - 1] = m \times U^k_i \setminus [2^{d+1} - 1] = T^k_{i,j} \setminus [2^{d+1} - 1].\end{aligned}$$

Hence $T^k_i \setminus [2^{d+1} - 1]$ is a $\Pi_2[+1, f]$-spectrum, and by Lemma 17, $T^k_i$ is a $\Pi_2[+1, f]$-spectrum. □

## 4.2 From spectra of Non-deterministic 2-counter automata to $\mathsf{FO}[+1, f]$-spectra

The following theorem is similar to Theorem 38, for the logic $\mathsf{FO}[+1, f]$.

**Theorem 53.** *For each non-deterministic 2-counter automaton $\mathcal{A}$, there exists a formula $\chi_{\mathcal{A}} \in \Pi_2[+1, f]$ such that $\mathsf{SP}(\chi_{\mathcal{A}}) = \left\{2^{t+3} + 4 \mid t \in \mathsf{SP}(\mathcal{A})\right\}$.*

*Proof.* Let $\mathcal{A}$ be a non-deterministic 2-counter automaton. We use Theorem 38 with $g(x) = x \times 2$, hence with $s^g_t = 2^{t+1}$. By this theorem, there exists $\varphi_{\mathcal{A}} \in \exists\mathsf{MSO}\Pi_2[+1, \times 2]$ such that $\mathsf{SP}(\varphi_{\mathcal{A}}) = \{2^{t+1} + 1 \mid t \in \mathsf{SP}(\mathcal{A})\}$. The formula $\varphi_{\mathcal{A}}$ is of the form $\exists X_0, \ldots, X_{d-1}.\psi_{\mathcal{A}}$ with $\psi_{\mathcal{A}} \in \Pi_2[+1, \times 2, X_0, \ldots, X_{d-1}]$. By Lemma 16, $\mathsf{SP}(\varphi_{\mathcal{A}}) = \mathsf{SP}(\psi_{\mathcal{A}})$.

Let $\chi_{\mathcal{A}} = \theta_{4,d,0} \wedge \Phi_{4,d,0}(\psi_{\mathcal{A}})$, where $\theta_{4,d,0}$ is the $\Pi_1[+1, f]$-formula of Lemma 51, and $\Phi_{4,d,0}$ is the function of Definition 48. By Lemma 52 we have:

$$\begin{aligned}\mathsf{SP}(\chi_{\mathcal{A}}) = \{4 \times n \mid n \in \mathsf{SP}(\psi_{\mathcal{A}})\} \setminus \left[2^{d+1} - 1\right] &= \left\{4 \times \left(2^{t+1} + 1\right) \mid t \in \mathsf{SP}(\mathcal{A})\right\} \setminus \left[2^{d+1} - 1\right] \\ &= \left\{2^{t+3} + 4 \mid t \in \mathsf{SP}(\mathcal{A})\right\} \setminus \left[2^{d+1} - 1\right].\end{aligned}$$

By Lemma 17, $\left\{2^{t+3} + 4 \mid t \in \mathsf{SP}(\mathcal{A})\right\}$ is also a spectrum. □

Using the ideas of the reduction from Theorem 41 to Theorem 53, we obtain the following theorem:

**Theorem 54.** *The finite satisfiability and the satisfiability over $\mathbb{N}$ of $\Pi_2[+1, f]$ are undecidable.*

## 4.3 Defining increasing functions in $\mathsf{FO}[+1, f]$

In last section, it is proven that $\Pi_2[+1, f]$ – with $f$ an uninterpreted unary function – allows to express in $\mathbb{N}$, that the interpretation of $f$ is the multiplication by a constant $b$. We now prove that we can encode functions that increase extremely quickly such as $n \mapsto c \uparrow^d n$, the Knuth's up-arrow function defined in [17].

We introduce the induction operator $\iota : ((\mathbb{N} \to \mathbb{N}) \times \mathbb{N}) \to (\mathbb{N} \to \mathbb{N})$. For $f : \mathbb{N} \to \mathbb{N}$, $c, n \in \mathbb{N}$, let $\iota(f, c)$ be the function defined by recursion on $c$ such that $\iota(f, c)(0) = c$ and $\iota(f, c)(n+1) = f(\iota(f, c)(n))$.

**Definition 55** ($\mathbb{F}$)**.** Let $\mathbb{F}$ be the smallest set of functions containing:

- the constant function 0,
- the function successor "$+1$",
- the function $f \circ g$ for each function $f$ and $g$ of $\mathbb{F}$ and
- the function $\iota(f, c)$ for each function $f$ of $\mathbb{F}$ and $c \in \mathbb{N}$.

We give some examples of functions in $\mathbb{F}$:

**Example 56.**

- the function $+1 \circ +(c-1)$ equals the function $+c$,
- the function $+c \circ 0$ equals the constant function $c$,
- the function $\iota(+c, 1)$ equals the multiplication by a constant $\times c$,
- the function $\iota(\times c, 1)$ equals the exponentiation in base $c$, $n \mapsto c^n$ and
- the function $c \uparrow^{d+1} n = \iota\left(c \uparrow^d n, c\right)$ equals the function $c \uparrow^d n$.

Hence, $\mathbb{F}$ contains functions that are known to increase quickly. On the other hand, it does not seem to contain functions such as $\log(n)$ or $n^2$.

**Theorem 57.** *For each $g \in \mathbb{F}$ there exists a formula $\varphi(x, y) \in \Pi_2[+1, f]$ such that for any $\{+1, f\}$-structure over the universe $\mathbb{N}$, $\mathcal{S} \models \varphi(x, y)$ if and only if $g\left(x^{\mathcal{S}}\right) = y^{\mathcal{S}}$.*

*Proof.* The function $g$ is defined by induction, using a finite number $m$ of functions. Let $(g_i)_{i \in [m-1]}$ be a list of those functions, with $g = g_m$. Let $M = m + 2$. Each function $g_i$ will be defined in the equivalence class $i + 2$ modulo $M$. Hence the set of multiples of $M$ must also be defined. For technical reasons, the last equivalence class defines the multiplication by $M$.

We state that the function $f^{\mathcal{S}}$ is such that for $n \in \mathbb{N}$, $f^{\mathcal{S}}(M \times n) = 0$, and for $a \in [M-1] \backslash \{0\}$, $f^{\mathcal{S}}(M \times n + a) \neq 0$. This can be express by the formula $\varphi_{\text{last}}$, equal to:

$$\bigwedge_{a=1}^{M-2} f(a) \neq 0 \wedge f(0) \doteq 0 \wedge \forall x. f(x) \doteq 0 \Longleftrightarrow f(x + M) \doteq 0.$$

By this requirement, $x \equiv a \mod (m+2)$ can be expressed by $f(x - a) \doteq 0$.

We state that $f^{\mathcal{S}}(n \times M + 1) = M^2 \times n + 1$, which can be express by the following formula $\varphi_{mul}$, equal to:

$$f(1) \doteq 1 \wedge (\forall x. x \equiv 1 \mod M)\, f(x + M) \doteq f(x) + M^2.$$

Hence $x \times M \doteq y$ can be expressed by:

$$\bigvee_{a=0}^{m+1} (x \equiv a \mod M \wedge f(x + 1 - a) + a \times M - 1 \doteq y),$$

Finally, for $a \in [m-1]$, for $n \in \mathbb{N}$, $f(n \times M + a + 2) = g_a(n) + 1$. Hence $g_a(x)$ can be defined as $f(x \times M + a + 2) - 1$ by the formula $\varphi_{g_a}$ defined as follows:

- if $g_a$ is the constant function 0, let $\varphi_{g_a}$ be $\forall x. f(x \times M + a + 2) \doteq 1$,
- if $g_a$ is the function $+1$, let $\varphi_{g_a}$ be $\forall x. f(x \times M + a + 2) \doteq x + 2$,

- if $g_a$ is the function $g_b \circ g_c$, let $\varphi_{g_a}$ be $\forall x. f\left(x \times M + a + 2\right) \doteq g_b\left(g_c(x)\right)$,
- and finally, if $g_a$ is the function $\iota\left(g_b, c\right)$, let $\varphi_{g_a}$ be $f\left(a\right) \doteq c + 1 \wedge \{\forall x. x \equiv a \mod M\}\, f\left(x + m + 2\right) \doteq g_b\left(f(x) - 1\right)$.

Then the formula $\varphi$ defined as

$$\varphi_{\text{last}} \wedge \varphi_{mul} \wedge \bigwedge_{i=0}^{m-1} \varphi_{g_i}.$$

asserts that $g$ is interpreted as required above.

Hence $g\left(x^{\mathcal{S}}\right) = y^{\mathcal{S}}$ can be expressed as $\varphi \wedge g_m(x) \doteq y$, that is as $\varphi \wedge f(Mx + m + 2) - 1 \doteq y$. □

# 5 Conclusion

In this paper, we have proven that, contrary to what may intuitively be expected, first-order logic with only the successor and an uninterpreted function is quite expressive. In particular the spectra of those logics are complex, even when there is only one alternation of quantifiers.

Note that the spectrum of $\varphi_{\mathcal{A}}$ for the non-deterministic 2-counter automaton $\mathcal{A}$ of Example 28 is equal to

$$\left\{2^{7n+3n^2+1} + 1 \mid n \in \mathbb{N}^{>0}\right\}.$$

We could generate spectra more natural than $\left\{2^{7n+3n^2+1} + 1 \mid n \in \mathbb{N}^{>0}\right\}$ such as $\left\{2^{n^2} + 1 \mid n \in \mathbb{N}^{>0}\right\}$. Indeed, the encoding of spectra used in Theorem 38 is robust. In order to give more flexibility to generate the desired spectra, the following modifications in the definition of the non-deterministic 2-counter automaton could easily be encoded in a similar fashion:

- adding any finite number of counters,
- adding an instruction that copies a counter into another one,
- assigning 0 to a counter.

More generally, it remains to find more classes of sets which are spectra of those logics. In particular, we are currently studying how to encode the image of polynomials as $\mathsf{FO}[+1, f]$-spectra with $f$ uninterpreted.

**Acknowledgments** We thank the anonymous referee for his/her remarks and suggestions to improve the paper. We thank Alexis Bès, our PhD advisor, for his help during the preparation of this paper. We thank the organizers of the Highlight Conference, where this result was first publicly presented. Finally, we thank Charles Paperman who asked at this conference whether a previous version of the result holds with two-variable logic.

# References

[1] E. Börger, E. Grädel, and Y. Gurevich, The classical decision problem, Perspectives in mathematical logic (Springer, Berlin, Heidelberg, New York, 1997).

[2] K. Gödel, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, Monatshefte für Mathematik und Physik **38**(1), 173–198 (1931).

[3] Y. Gurevich, The decision problem for standard classes, J. Symb. Log. **41**(2), 460–464 (1976).

[4] A. Ehrenfeucht, Decidability of the theory of the linear ordering relation, Notices Amer. Math. Soc. **6(3):268**, 556–38 (1959).

[5] A. Durand, N. D. Jones, J. A. Makowsky, and M. More, Fifty years of the spectrum problem: survey and new results, Bulletin of Symbolic Logic **18**(12), 505–553 (2012).

[6] A. Durand, R. Fagin, and B. Loescher, Spectra with only unary function symbols, in: Computer Science Logic, edited by M. Nielsen and W. Thomas, Lecture Notes in Computer Science Vol. 1414 (Springer Berlin Heidelberg, 1998), pp. 189–202.

[7] Y. Gurevich and S. Shelah, Spectra of monadic second-order formulas with one unary function, in: 18th IEEE Symposium on Logic in Computer Science (LICS 2003), 22-25 June 2003, Ottawa, Canada, Proceedings, (2003), pp. 291–300.

[8] M. Presburger, Über die vollständigkeit eines gewissen systems der arithmetik ganzer zahlen, in welchen, die addition als einzige operation hervortritt, in: Comptes Rendus du Premier Congrès des Mathématicienes des Pays Slaves, (Warsaw, 1927), pp. 92–101, 395.

[9] H. Straubing, Finite Automata, Formal Logic, and Circuit Complexity (Birkhäuser, 1994).

[10] P. Chocron, P. Fontaine, and C. Ringeissen, A gentle non-disjoint combination of satisfiability procedures, in: Automated Reasoning - 7th International Joint Conference, IJCAR 2014, Held as Part of the Vienna Summer of Logic, VSL 2014, Vienna, Austria, July 19-22, 2014. Proceedings, (2014), pp. 122–136.

[11] É. Ailloud and A. Durand, The expressive power of bijections over weakly arithmetized structures, Theory Comput. Syst. **39**(2), 297–309 (2006).

[12] R. E. Shostak, A practical decision procedure for arithmetic with function symbols, J. ACM **26**(2), 351–360 (1979).

[13] W. Thomas, A note on undecidable extensions of monadic second order successor arithmetic, Archiv für mathematische Logik und Grundlagenforschung **17**, 43–44 (1975).

[14] A. Cobham, On the base-dependence of sets of numbers recognizable by finite automata, Mathematical systems theory **3**(2), 186–192 (1969).

[15] V. Bruyère, G. Hansel, C. Michaux, and R. Villemaire, Logic and p-recognizable sets of integers, Bull. Belg. Math. Soc **1**, 191–238 (1994).

[16] A. Milchior, Undecidability of satisfiability of expansions of FO[<] with a semilinear non regular predicate over words., Computability in Europe, Informal Proceedings (2013).

[17] D. E. Knuth, Mathematics and Computer Science: Coping with Finiteness, STAN-CS- (Department of Computer Science, Stanford University, 1976).

[18] H. D. Ebbinghaus, J. Flum, and W. Thomas, Mathematical logic (2. ed.), Undergraduate texts in mathematics (Springer, 1994).

[19] M. L. Minsky, Recursive Unsolvability of Post's Problem of "tag": And Other Topics in Theory of Turing Machines, Group report (Massachusetts Institute of Technology, Lincoln Laboratory, 1960).