CDMTCS
Research
Report
Series

Definability and Regularity in Automatic Presentations of Subsystems of Arithmetic

Bakhadyr Khoussainov
Department of Computer Science
University of Auckland, New Zealand

Sasha Rubin
Department of Mathematics
University of Auckland, New Zealand

Frank Stephan
Mathematical Institute
University of Heidelberg, Germany

CDMTCS-209
January 2003

Centre for Discrete Mathematics and
Theoretical Computer Science

Definability and Regularity in Automatic Presentations of Subsystems of Arithmetic

Bakhadyr Khoussainov[1] and Sasha Rubin[2] and Frank Stephan[3]

[1] Computer Science Department, The University of Auckland, New Zealand
bmk@cs.auckland.ac.nz
[2] Mathematics Department, The University of Auckland, New Zealand
rubin@math.auckland.ac.nz
[3] Mathematical Institute, University of Heidelberg, Germany
fstephan@math.uni-heidelberg.de

Abstract. This paper is devoted to the study of the relationship between regularity and definability of relations in structures presented by finite automata. An emphasis is given to relations in fragments of arithmetic, in particular in $(\omega, \leq)$, (ω, S), $(\omega, +)$ and some of their variants. A relation in a structure is *intrinsically regular* if it is regular in every automatic presentation of the structure. All definable relations in the first order logic with finite number of parameters are intrinsically regular. We investigate questions related to whether or not intrinsically regular relations are definable. For example, on the one hand, we show that the set M^2 of all even numbers in every automatic presentation of $(\omega, \leq)$ is intrinsically regular (but not definable). On the other, we show that there exists an automatic presentation of (ω, S) in which the set M^2 is not regular. In particular, we show that a unary relation in (ω, S) is intrinsically regular if and only if it is definable.

1 Introduction

The aim of this paper is to study the relationship between regularity and definability in automatic structures. This is an old topic that goes back to the results of Büchi in the 1960's who proved, for example, that a language L is regular if and only if it is definable in the monadic second order logic over finite words. This is closely related to work in which regular languages are characterised via definability in certain fragments of arithmetic. For example, work of Büchi, Cobham, Semenov, Muchnik, Bruyère et al. (see [2]) a for good exposition), is devoted to characterising those relations in fragments of arithmetic that can be recognised by finite automata. In this paper we continue this line of research in light of current interest and results in the theory of automatic structures. We are also motivated by the intense investigations in computable model theory on the relationship between computability and definability, a research program first proposed by Ash and Nerode in their famous paper on intrinsically recursive relations [1]. In this paper we refine the Ash-Nerode program by considering the class of automatic structures rather than the much general class of computable structures.

Assume, we have a structure $\mathcal{A}$ that is automatically presentable (Definition 23). This means that there is an encoding of the domain A of $\mathcal{A}$ such that the domain, functions and relations of $\mathcal{A}$ are regular. Let $R \subset A^m$ be a relation in $\mathcal{A}$, though not necessarily in the language of $\mathcal{A}$. For example, R can be the reachability relation if $\mathcal{A}$ is a graph, or R can be the dependency relation if $\mathcal{A}$ is a group. Automata theoretic properties of R depend on automatic presentations of $\mathcal{A}$. For example, in one presentation of $\mathcal{A}$ the relation R may be recognised by an n state automaton, in another presentation R may be recognised by an m state automaton with $n \neq m$, and there could be automatic presentations of $\mathcal{A}$ in which R is not regular. One of our goals is to study those relations in $\mathcal{A}$ that are regular under every automatic presentation of $\mathcal{A}$. We single out such relations in the following definition.

Definition 11 (cf. [1]) *A relation R in an automatic structure $\mathcal{A}$ is* **intrinsically regular** *if for every automatic structure $\mathcal{B}$ isomorphic to $\mathcal{A}$ the image of the relation R in $\mathcal{B}$ is regular.*

Thus the intrinsically regular relations in $\mathcal{A}$ are the ones for which regularity is invariant under all automatic presentations of $\mathcal{A}$. A natural class of intrinsically regular relations is the class of all definable relations. We say that a relation R is **definable** in $\mathcal{A}$ if there exists a formula $\phi(\bar{x}, \bar{c})$ in the language of the first order logic extended by the $\exists^\infty$ (there exist infinitely many) quantifier with finitely many parameters $\bar{c}$ from $\mathcal{A}$ such that $R = \{\bar{a} \mid \mathcal{A} \models \phi(\bar{a}, \bar{c})\}$. Definable relations in a given automatic structure are intrinsically regular as provided by the next theorem (see [7] or [3]).

Theorem 12 *Given an automatic structure $\mathcal{A}$ and a relation R that is definable in $\mathcal{A}$, one can effectively construct an automaton recognising R.* □

In light of this, we name the class of automatic structures for which the intrinsic regular relations are exactly the definable ones.

Definition 13 *An automatic structure $\mathcal{A}$ is* **intrinsically regular** *if every intrinsically regular relation R in $\mathcal{A}$ is definable.*

We illustrate these concepts with a well studied example, the standard model of Presburger Arithmetic, that is the structure $(\omega, +)$. For each $m \in \omega$ consider the presentation $\mathcal{A}_m$ of ω over the alphabet $\Sigma_m = \{0, \dots, m-1\}$. Here the natural number $n \in \omega$ is coded in A_m as the reverse of its base m-representation. For $R \subseteq \omega^n$, write R_m for the image of R in A_m. Then for $m > 1$, $(A_m, +_m)$ is automatic (Definition 23) because an automaton for $+_m$ can essentially implement the standard algorithm for base m addition. The Cobham-Semenov theorem (see [2]) states that for $R \subseteq \omega^n$, if (A_k, R_k) and (A_l, R_l) are both automatic structures for multiplicatively independent k and l, then R is definable in $(\omega, +)$. Here k and l are multiplicatively independent if $(\forall n, m > 0) k^n \neq l^m$. So if R is intrinsically regular for $(\omega, +)$ then it is definable in $(\omega, +)$. Thus, the Cobham-Semenov implies that:

Theorem 14 *The structure $(\omega, +)$ is intrinsically regular.* □

In this paper we concentrate on subsystems of the weak arithmetic $(\omega, S, +, \leq)$. In particular, we concentrate on the relationship between automatic presentations and intrinsically regular relations in the structures $(\omega, +)$, $(\omega, \leq)$, and

(ω, S) and among the binary relation $\leq$ and the unary relations $M^k = \{n \in \omega \mid (\exists j \in \omega) n = kj\}$ and $P^k = \{n \in \omega \mid (\exists j \in \omega) n = k^j\}$. We also investigate how to expand these structures in order to make them intrinsically regular.

Here is now an outline of the paper. The next section gives necessary definitions and some examples. In Section 3, we study the structure $(\omega, +)$ and show that in *every* automatic presentation of $(\omega, +)$ there is a regular relation in it whose image is not regular in another automatic presentation of $(\omega, +)$. Section 4 investigates regular relations in the structure $(\omega, \leq)$. We show that $(\omega, \leq)$ is not intrinsically regular. For example, we prove that the set M^2 of all $n \in \omega$ at even positions is an intrinsically regular unary relation but M^2 is not definable. Section 5 studies intrinsic regularity of the structure (ω, S). The main result of the section is that a unary relation R is intrinsically regular if and only if it is finite or co-cofinite. Finally, we briefly study trees and investigate the relationship between regularity of the partial order and successor relation on trees.

2 Automata Preliminaries

A thorough introduction to automatic structures can be found in [3] and [7]. A recent survey paper [8] discusses the basic results and possible directions for future work in the area. Familiarity with the basics of finite automata theory is assumed though for completeness the necessary definitions are included here.
A *finite automaton* $\mathcal{A}$ over an alphabet Σ is a tuple (S, ι, Δ, F), where S is a finite set of *states*, $\iota \in S$ is the *initial state*, $\Delta \subset S \times \Sigma \times S$ is the *transition table* and $F \subset S$ is the set of *final states*. A *computation* of $\mathcal{A}$ on a word $\sigma_1\sigma_2 \ldots \sigma_n$ ($\sigma_i \in \Sigma$) is a sequence of states, say $q_0, q_1, \ldots, q_n$, such that $q_0 = \iota$ and $(q_i, \sigma_{i+1}, q_{i+1}) \in \Delta$ for all $i \in \{0, 1, \ldots, n-1\}$. If $q_n \in F$, then the computation is *successful* and we say that automaton $\mathcal{A}$ *accepts* the word. The *language* accepted by the automaton $\mathcal{A}$ is the set of all words accepted by $\mathcal{A}$. In general, $D \subset \Sigma^\star$ is *finite automaton recognisable*, or *regular*, if D is equal to the language accepted by $\mathcal{A}$ for some finite automaton $\mathcal{A}$.

Classically finite automata recognise sets of words. The following definitions extends recognisability to relations of arity n, called *synchronous n–tape automata.* Informally a synchronous n–tape automaton can be thought of as a one-way Turing machine with n input tapes [4]. Each tape is regarded as semi-infinite having written on it a word in the alphabet Σ followed by an infinite succession of blanks, $\diamond$ symbols. The automaton starts in the initial state, reads simultaneously the first symbol of each tape, changes state, reads simultaneously the second symbol of each tape, changes state, etc., until it reads a blank on each tape. The automaton then stops and accepts the n–tuple of words if it is in a final state. The set of all n–tuples accepted by the automaton is the relation recognised by the automaton. Here is a formalisation:

Definition 21 *Let $\Sigma_\diamond$ be $\Sigma \cup \{\diamond\}$ where $\diamond \notin \Sigma$. The* convolution of a tuple *$(w_1, \cdots, w_n) \in (\Sigma^*)^n$ is the tuple $(w_1, \cdots, w_n)^\diamond \in ((\Sigma_\diamond)^*)^n$ formed by concatenating the least number of blank symbols, $\diamond$, to the right ends of the w_i,*

$1 \leq i \leq n$, *so that the resulting words have equal length. The* convolution of a relation $R \subset (\Sigma^*)^n$ *is the relation* $R^\diamond \subset ((\Sigma_\diamond)^*)^n$ *formed as the set of convolutions of all the tuples in* R.

Definition 22 An n–tape automaton *on* Σ *is a finite automaton over the alphabet* $(\Sigma_\diamond)^n$. *An* n*–ary relation* $R \subset \Sigma^{\star n}$ *is* finite automaton recognisable *or* regular *if its convolution* $R^\diamond$ *is recognisable by an* n*–tape automaton.*

We now relate n–tape automata to structures. A *structure* $\mathcal{A}$ consists of a set A called the *domain* and some constants, relations and operations on A. We may assume that $\mathcal{A}$ only contains relational and constant predicates as the operations can be replaced with their graphs. We write $\mathcal{A} = (A, R_1^A, \ldots, R_k^A, c_0^A, \ldots, c_t^A)$ where R_i^A is an n_i–ary relation on $\mathcal{A}$ and c_j^A is a constant element of $\mathcal{A}$.

Definition 23 *A structure* $\mathcal{A}$ *is* automatic *over* Σ *if its domain* $A \subset \Sigma^\star$ *and the relations* $R_i^A \subset \Sigma^{*n_i}$ *all are finite automaton recognisable. An isomorphism from a structure* $\mathcal{B}$ *to an automatic structure* $\mathcal{A}$ *is an* automatic presentation *of* $\mathcal{B}$ *in which case* $\mathcal{B}$ *is called* automatically presentable (*over* Σ). *A structure is called* automatic *if it is automatic over some alphabet.*

Examples of automatically presentable structures are the fragments of arithmetic considered here, the group of integers $(Z, +)$, the order on the rationals $(Q, \leq)$, the Boolean algebra of finite or co-finite subsets of ω. An example of an automatic structure is the word structure $(\{0,1\}^\star, L, R, E, \preceq)$, where for all strings $x, y \in \{0,1\}^\star$ we have $L(x) = x0$, $R(x) = x1$, $E(x, y)$ iff $|x| = |y|$, and $\preceq$ is the lexicographical order.

So an automatic structure is one that is explicitly given by finite automata that recognise the domain and the basic relations of the structure. An automatically presentable structure is one that is isomorphic to some automatic structure. Informally, automatically presentable structures are those that have finite automata implementations. This, the same structure may have different (indeed infinitely many) automatic presentations. It is a goal of this paper to study the relationships between different automatic presentations of a given structure, and understand how queries about the this structure change (automata theoretically) when one varies the presentation.

3 Closed-regular structures

One can tighten the notion of intrinsic regular structure to require that regularity and definability coincide. It turns out that there are natural examples of such automatic structures. We first give a definition:

Definition 31 *An automatic structure* $\mathcal{A}$ *is* **closed-regular** *if every regular relation in* $\mathcal{A}$ *is definable in* $\mathcal{A}$.

Note that if a structure $\mathcal{A}$ has a closed-regular presentation then $\mathcal{A}$ is intrinsically regular. Indeed, assume that $\mathcal{A}$ has a closed-regular (hence automatic) presentation $\mathcal{B}$. Let R be an intrinsically regular relation in $\mathcal{A}$. Then R is regular in $\mathcal{B}$. Therefore R is definable. However, we note the notion of closed-regular is sensitive in the sense that if $\mathcal{A}$ and $\mathcal{B}$ are both automatic presentations of the same structure, $\mathcal{A}$ may be closed-regular while $\mathcal{B}$ may not be. We illustrate these concepts with some natural examples. Consider the automatic structure $\mathcal{U}_p = (A_p, +, |_p)$ for $p > 1$, where A_m is an automatic presentation of ω over the alphabet $\Sigma_m = \{0, \ldots, m-1\}$, and $x|_p y$ if x is a power of p and x divides y. A relation $R \subset A_p^n$ is regular if and only if it is definable in $\mathcal{U}_p$ see [2] or [3]. So $\mathcal{U}_p$ is closed-regular for $p > 1$. Similarly, the automatic presentation $\mathcal{A} = (1^*, \leq, (M^k)_{k \in \omega})$ of $(\omega, \leq, (M^k)_k)$, is closed-regular since it is complete for the class of unary automatic structures (see [3]). However this structure has an automatic presentation over the binary strings $\mathcal{B} = (A_2, \leq, (M^k)_{k \in \omega})$ that is not closed-regular, since for example $+$ is regular in $\mathcal{B}$. Note that in the last example, we are allowing the signature to be infinite.

We have already seen that $(\omega, +)$ is intrinsically regular. It is also an example of a structure that does not have a closed-regular presentation. Recall that a subset M of ω is *eventually periodic* if there exists $n_0, p \in \omega$ such that for all $x > n_0$, $x \in M$ if and only if $x + p \in M$.

Lemma 32 [5, Theorem 32F] *A unary relation is definable in* $(\omega, +)$ *if and only if it is an eventually periodic set.*

Proof. The structure $(\omega, +, (x \equiv_m y)_{m \in \omega})$ admits quantifier elimination and the eventually periodic subsets of ω are closed under union and complementation. □

Theorem 33 *The structure* $(\omega, +)$ *does not have a closed-regular presentation.*

Proof. Assume that (D, E) is an automatic presentation of $(\omega, +)$. We construct a unary regular relation R in (D, E) that is not definable in $(\omega, +)$. Without loss of generality, the empty string is in D. There is a constant c such that for every $x \in D$, the length of $x + x$ is at most $|x| + c$. For every n let

$$D_n = \{x \in D \mid |x| \leq c \cdot n\};$$
$$R = \{x \in D \mid \exists n (x \in D_n \wedge x = \max(D_n))\};$$

where in the definition of R the maximum refers to the ordering of $\leq$ of $(\omega, +)$. The relation R is regular in the given presentation because the binary relation $D' = \{(x, y) \mid x \in D,\ |x| \leq |y|, \text{ and } |y| \text{ is a multiple of } c\}$ is regular.

If $x, y \in R$ and $x < y$, then x, y belong to different sets D_n, D_m, respectively with $n < m$. Since $x + x$ has length at most $|x| + c$, $x + x \in D_m$ for all $m > n$. So if $x \in R$ then the next number in R larger than x is at least $x + x$. So R is not eventually periodic and so by Lemma 32, R is not definable in $(\omega, +)$. □

Corollary 34 *Every automatic presentation of* $(\omega, +)$ *has a regular set that is not intrinsically regular.*

Proof. Consider the presentation (D, E) and the regular but not definable relation R above. By Theorem 14, R can not be intrinsically regular. □

We end this section with some basic properties of intrinsically regular relations:

Proposition 35 *The intrinsically regular relations in an automatic structure $\mathcal{A}$ are closed under union, complementation and projection. Moreover, a non intrinsically regular relation R in $\mathcal{A}$ is not intrinsically in every reduct $\mathcal{B}$ of $\mathcal{A}$.*

Proof. For the first suppose R_0, R_1 are n–ary intrinsically regular relations in $\mathcal{A}$. Then for every automatic presentation $\mathcal{B}$ of $\mathcal{A}$, the structure $(\mathcal{B}, R_0, R_1)$ is automatic. Since regular relations are closed under finite union, complementation and projection, we have $R_0 \cup R_1, A^n \setminus R$ and the projection $\pi(R)$ are regular in $\mathcal{B}$. This proves the first part. For the second part it suffices note that every automatic presentation of $\mathcal{A}$ is also an automatic presentation of $\mathcal{B}$. □

4 Intrinsic regularity in $(\omega, \leq)$

In this section we consider the linearly ordered structure $(\omega, \leq)$. We will show that $(\omega, \leq)$ has non definable but intrinsically regular relations. The following lemma describes the definable unary relations in $(\omega, \leq)$.

Lemma 41 [5, Theorem 32A] *A unary relation U is first-order definable in $(\omega, \leq)$ if and only if it is either finite or co-finite.*

Proof. Since $\exists^\infty x \Phi(x)$ is equivalent in $(\omega, \leq)$ to $\forall z \exists x (x > z \wedge \Phi(x))$, consider the usual first order theory of $(\omega, \leq)$. It admits quantifier elimination. Hence every definable set U in $(\omega, \leq)$ is either finite or co-finite. The converse is straightforward. □

It turns out that the structure $(\omega, \leq)$ is not intrinsically regular. Recall that $M^i \subseteq \omega$ for $i \in \omega$ is the set of all numbers $n \in \omega$ that are multiples of i. These sets are not definable in $(\omega, \leq)$.

Theorem 42 *The unary predicate M^2 is intrinsically regular for $(\omega, \leq)$.*

Proof. Let (D, E) be an automatic copy of $(\omega, \leq)$ over alphabet Σ and let $C \subseteq D$ be the words corresponding to the even natural numbers. Then $x \in C$ if and only if $\{y \in D \mid E(y, x)\}$ has odd cardinality. We will define an automaton over Σ accepting such x's. Let $\mathcal{A} = (Q_A, \iota, \Delta_A, F_A)$ be the automaton over Σ recognising E. The idea is that the new automaton calculates the parity of the number of paths in $\mathcal{A}$ with first component x and accepts x when the parity of the number of successful paths is odd.

Alter $\mathcal{A}$ so that instead of accepting only $(x, y)^\diamond$ for $x \leq y$, it also accepts $(x, y)^\diamond \binom{\diamond}{\diamond}^n$ for all $n \in \omega$ and every $x \leq y$. In other words, if $x = x_0 x_1 \ldots x_k \in D$ and $y = y_0 y_1 \ldots y_l \in D$ with say $k \leq l$, then $x \leq y$ if and only if $\mathcal{A}$ accepts exactly the words of the form $\binom{x_0}{y_0}\binom{x_1}{y_1} \ldots \binom{x_k}{y_k}\binom{\diamond}{y_{k+1}} \ldots \binom{\diamond}{y_l}\binom{\diamond}{\diamond}^s$ for all $s \in \omega$.

As a first step define an automaton $\mathcal{B}$ as $(Q_B, \iota_B, \Delta_B, F_B)$ over $\Sigma_\diamond$ as follows.

- Q_B is the set of all subsets of Q_A.

- $\Delta_B(V,\sigma) = W$ where W consists of all states $q \in Q_A$ such that the cardinality of $\{(v,\sigma') \mid q = \Delta_A(v, \binom{\sigma}{\sigma'}), v \in V \wedge \sigma' \in \Sigma_\diamond\}$ is odd.
- $\iota_B = \{\iota\}$.
- F_B consists of those states V such that $|V \cap F_A|$ is odd.

In words, $q \in \Delta_B(V,\sigma)$ means that there are an odd number of edges in $\mathcal{A}$ from states of V to the state q labeled with first component σ.

Now define P_y as the set of all paths in $\mathcal{A}$ starting at ι and labeled with y in the first component. Note that $y \in \Sigma^*\{\diamond\}^*$ and $\mathcal{B}$ only runs on words of this form.

Claim. For $|y| \neq 0$, $|P_y|$ is odd if and only if $\Delta_B(\iota_B, y) \neq \emptyset$.

Let $y \in \Sigma^*\{\diamond\}^*$ and let $V_0, V_1, \ldots, V_n, V_{n+1}, \ldots, V_{n+k}$ be a run in $\mathcal{B}$ on y. Proceed by induction on $|y|$. If $|y| = 1$ then by definition $q \in V_1$ if and only if there are an odd number of edges in $\mathcal{A}$ from ι to q if and only if $|P_y|$ is odd, as required. Suppose that $|y| = n$ and that the hypothesis holds for y. Consider y extended by σ. Then by definition, $q \in \Delta_B(V,\sigma)$ if and only if there are an odd number of edges from states of V to q with first component labeled σ.
Note that every path in $P_{y\sigma}$ extends a path in P_y by an edge. Now if $|P_y|$ were even then no matter how many extensions there are from a path in P_y to one in $P_{y\sigma}$, $|P_{y\sigma}|$ is also even. So assume that $|P_{y\sigma}|$ is odd. Then so is $|P_y|$ which implies that $V \neq \emptyset$ and for every $\tilde{q} \in V$ that there are an odd (and in particular at least 1) number of edges in $\mathcal{A}$ leaving $\tilde{q}$ labeled with first component σ. Hence $\Delta(V,\sigma) \neq \emptyset$. Conversely, suppose that $\Delta(V,\sigma) = \emptyset$. That is, for every $q \in Q_A$ there are an even number of edges in $\mathcal{A}$ from states of V to the state q labeled with first component σ. Hence $|P_{y\sigma}|$ is even if $V \neq \emptyset$. Alternatively if $V = \emptyset$ then by induction $|P_y|$ is even and so $|P_{y\sigma}|$ must also be even. This completes the proof of the claim.

Hence we immediately have that $\mathcal{B}$ accepts exactly those words $y \in \Sigma^*\{\diamond\}^*$ where the number of paths in $\mathcal{A}$ from ι labeled with first component y and ending in an accept state of $\mathcal{A}$ is odd. Now there exists a constant $\kappa \in \omega$ such that if t is the length of the longest word in $\{z \in D : E(z,y)\}$ then $t - |y| < \kappa$. For otherwise by the pumping lemma on $\mathcal{A}$ there would exists infinitely many words $\leq y$, contradicting that $\{z \in D : E(z,y)\}$ is finite. So write y' as y followed by κ many $\diamond$'s. Then $z \leq y$ if and only if $\mathcal{A}$ accepts $(z,y')^\diamond$. So y' is accepted by $\mathcal{B}$ if and only if there are an odd number of strings $z \in \Sigma^*\{\diamond\}^*$ of length $|y'|$ where (z,y') is accepted by $\mathcal{A}$. So finally construct an automaton $\mathcal{B}'$ that accepts y if and only if $\mathcal{B}$ accepts y'. Then $\mathcal{B}'$ accepts $x \in D$ if and only if the set $\{z \in D : E(z,x)\}$ has odd cardinality. This completes the proof. □

So combining the fact the M^2 is intrinsically regular for $(\omega, \leq)$ and is not definable in $(\omega, \leq)$ one has the following theorem as a corollary:

Theorem 43 *The structure $(\omega,\leq)$ is not intrinsically regular.* □

The next corollary describes all the intrinsically regular unary relations of $(\omega, \leq)$:

Corollary 44 *A unary relation R is intrinsically regular for $(\omega, \leq)$ if and only if R is a regular subset of ω.*

Proof. Adapting the proof above, yields that for each i the unary relation M^i is intrinsically regular. So suppose that $R \subseteq \omega$ is intrinsically regular for $(\omega, \leq)$. Then in particular $(A_1, \leq_1, R_1)$ is an automatic structure and so R_1 is a regular subset of $\{0\}^*$. Conversely, if R is a regular subset of ω then it is a finite union of arithmetic progressions. So R is a boolean combination of sets of the form M^i for $i \in \omega$. Each M^i is intrinsically regular for $(\omega, \leq)$ by Theorem 42, so is R. □

Proposition 45 *The ternary relation $+$ is not intrinsically regular for $(\omega, \leq)$.*

Proof. Define $D = \{0\}^*$ and $E(x, y)$ iff $|x| \leq |y|$. Then the relation $U(x, y) \Leftrightarrow x + x = y$ just says for $x, y \in \{0\}^*$ that y has the double length of x. A pumping lemma argument shows that this relation is not regular. □

5 Intrinsic Regularity in (ω, S)

Consider the structure (ω, S) where S is the successor function. Our goal is to show that in this structure, the intrinsically regular unary relations are those that are either finite or co–finite. For this we first not that as a consequence of Proposition 41, one has:

Corollary 51 *A definable unary relations in (ω, S) is finite or co-finite.* □

The following theorem is the main result of this section.

Theorem 52 *The relations M^2 and $\leq$ are not intrinsically regular for (ω, S).*

Proof. We construct an automatic structure $(\{0, 1\}^*, f)$ that is isomorphic to (ω, S). For constructing f, the following auxiliary notions are required: Given a string $x \in \{0, 1\}^*$, let $ev(x)$ be the binary number represented by the bits in the even positions of x and $od(x)$ be the binary number represented by the bits in the odd positions. Let n (m respectively) be the number of bits in the even (odd) positions of x; note that $m \leq n \leq m + 1$ and $|x| = m + n$. For example, if $x = 10101$ then $ev(x) = 111, od(x) = 00, n = 3$ and $m = 2$. The successor of x is computed according to the first case that applies in the following list:

1. If $n \leq 2$ then $f(x)$ is the successor of x with respect to length-lexicographic ordering.
2. If $ev(x) = 2^{n-1}$ and $od(x) = 2^m - 1$ then $f(x) = 0^{n+m+1}$.
3. If $ev(x) = 2^{n-1}$ and $od(x) < 2^m - 1$ then $f(x) = y$ for the y with $|y| = |x|$, $ev(y) = 0$ and $od(y) = od(x) + 1$.
4. If $ev(x) = 2^{n-1} - 2od(y) - 1$ modulo 2^n then $f(x) = y$ for the y with $|y| = |x|$, $ev(y) = ev(x) + 4od(y) + 2$ and $od(y) = od(x)$.
5. If $ev(x) = 2^n - 2od(y) - 1$ modulo 2^n then $f(x) = y$ for the y with $|y| = |x|$, $ev(y) = 2^n - 1$ and $od(y) = od(x)$.
6. Otherwise $f(y) = y$ for the y with $|y| = |x|$, $ev(y) = ev(x) + 2od(y) + 1$ and $od(y) = od(x)$.

The operations given here are automatic since the addition is automatic, the i'th bit of $ev(x)$ and of $od(x)$ are adjacent in x, and all constants are easily recognizable: 2^{n-1} is represented by a 1 and $n-1$ 0's following it, 2^n is represented by n 0's and $2^m - 1$ is represented by m 1's.

Note that for every natural number k and odd number h there is a number $u \in \{0, 1, \ldots, 2^n-1\}$ such that $k = h \cdot u$ modulo 2^n. Let π be an isomorphism from $(\{0,1\}^*, f)$ to (ω, S). One can inductively verify the existence of the isomorphism and in particular show the following for the case that $n \geq 3$.

- There is a unique string x' such that $|x'| = |x|$, $ev(x') = 0$ and $od(x') = od(x)$.
- There is a unique number $u \in \{0, 1, \ldots, 2^n-1\}$ such that $ev(x) = u \cdot (2od(x)+1)$ modulo 2^n.
- If $u < 2^{n-1}$ then $\pi(x) = \pi(x') + u$.
- If $u = 2^{n-1}$ then $\pi(x) = \pi(x') + 2^n - 1$.
- If $u > 2^{n-1}$ then $\pi(x) = \pi(x') + u - 1$.
- If $ev(y) = 0$ and $od(y) = od(x') + 1 \leq 2^m - 1$ then $\pi(y) = \pi(x) + 2^n$.

¿From this one can conclude that the successor always goes through all elements of length n and then goes to length $n+1$. Thus $(\{0,1\}^*, f)$ is an automatic representation of (ω, S).

Recall that $M^2 = \{x \mid \pi(x) \text{ is odd}\}$. Note that $\pi(x')$ is odd for all non-empty strings x. Thus one has that $x \in M^2$ iff either $u < 2^{n-1}$ and $ev(x)$ is even or $u \geq 2^{n-1}$ and $ev(x)$ is odd. So knowing whether $x \in M^2$ and $ev(x)$ is even permits to decide whether $u \geq 2^{n-1}$. The goal of the construction was to make this impossible.

To show that the goal is reached, assume the following special case: The number n is odd, x has n 1's in the even positions and $m - r$ 0's followed by r 1's in the odd position. Furthermore, assume that $n \leq 2r + 3$ and $r + 1 < n$. Now it is shown that under these premises the membership of x in M^2 is equivalent to $n \neq 2r + 3$ – a condition that cannot be checked by a finite automaton.

If $n = 2r + 3$ then $2^n - 1 = (2^{n-1} + 2^{r+1} + 1) \cdot (2^{r+1} - 1)$ modulo 2^n. It follows that $\pi(x) = \pi(x') + 2^{n-1} + 2^{r+1}$ and $x \notin M^2$.

If $n < 2r + 3$ then $2^n - 1 = (2^{r+1} + 1) \cdot (2^{r+1} - 1)$ modulo 2^n. It follows that $\pi(x) = \pi(x') + 2^{r+1} + 1$ and $x \in M^2$. So M^2 is not regular.

Theorem 42 showed that every automatic presentation of $(\omega, \leq)$ satisfies that M^2 is regular. Thus $\leq$ is not regular in the structure $(\{0,1\}^*, f)$. □

The next result shows that there are also automatic presentations of (ω, S) where M^2 is regular but $\leq$ is not.

Theorem 53 *The ordering $\leq$ is not intrinsically regular for $(\omega, S, M^2, M^3, \ldots)$.*

Proof. We construct a presentation (D, g) over $\Sigma = \{0, 1\}$ of (ω, S) in which all sets M^k are regular, but $\leq$ is not. Let $D = 0^*1^*$. Now define

$$g(0^n1^m) = \begin{cases} 0^n1 & \text{if } m + 2n \text{ is even and } m = 0; \\ 0^{n+1}1^{m-2} & \text{if } m + 2n \text{ is even and } m > 0; \\ 1^{m+1} & \text{if } m + 2n \text{ is odd and } n = 0; \\ 0^{n-1}1^{m+2} & \text{if } m + 2n \text{ is odd and } n > 0. \end{cases}$$

It is easy to see that g is automatic. Furthermore, g corresponds to an ordering where $0^n1^m \prec 0^{n'}1^{m'}$ whenever $m+2n < m'+2n'$. The mapping $0^n1^m \to m+2n$ might be viewed as a norm of the strings in D compatible with g. This table with the first strings ordered by g and with the corresponding norm given

norm	values of this norm ordered by g
0	λ
1	1
2	11, 0
3	01, 111
4	1111, 011, 00
5	001, 0111, 11111
6	111111, 01111, 0011, 000

shows how the ordering works. Let $\preceq$ be the ordering induced by g. Then $1^m \preceq 0^n$ iff $m \leq 2n$. Such a condition is not regular and thus $\preceq$ is not automatic.

Let π be the isomorphism from (D, g) to (ω, S). The lengths of x and $g(x)$ differ by 1. Thus, $\pi(0^n1^m)$ is odd iff the length $n+m$ of the string is odd. So, M^2 is represented by a regular set. For general M^k, note that the value of $\pi(0^n1^m)$ is determined by the value of the norm $2n+m$ modulo $2k$ and $n+m$ modulo k. This follows from the following facts: there are exactly $h(h+1)$ many strings of norm strictly below h; if the norm is even then g decreases the length in every step; if the norm is odd then g increases the length in every step. So a finite automaton can compute whether 0^n1^m is in M^k by determining the values of $2n+m$ modulo $2k$, $n+m$ modulo k and then looking up the result in a table with $2k \cdot k$ many entries. □

6 Intrinsic regularity in structures between (ω, S) and $(\omega, \leq)$

Given a partial ordering $(T, \preceq)$, let S be the successor relation in $(T, \preceq)$ where $S(x, y)$ iff $x \prec y$ and there is no z with $x \prec z \prec y$. In particular, in the structure $(\omega, \leq)$ the notion $S(x, y)$ is equivalent to $y = x+1$. Furthermore, define $P^k \subseteq N$ as the set of powers of k. Write $\mathcal{P}^2$ for the structure (ω, f, P^2) where f is the successor function. In the proof of Theorem 53 the set P^2 is regular in the constructed presentation (D, g). So we strengthen this to:

Proposition 61 *1. The ordering $\leq$ is not intrinsically regular for $\mathcal{P}^2$.*
2. The set M^i is not intrinsically regular for $\mathcal{P}^2$, for every $i > 1$.

The next example shows that $(T, \preceq)$ is a partial ordered structure such that (T, S) has an automatic copy but $(T, \preceq)$ does not. The example will be a tree where each node has one or two successors, so it will be sufficiently natural. The proof of the result uses the notion of the Cantor-Bendixson rank for trees that is defined as follows.

Definition 62 [6] *Given a tree $\mathcal{T}$, for every countable ordinal α the α's derivative $d^\alpha(\mathcal{T})$ is defined by transfinite induction as follows.*

- $d^0(\mathcal{T})$ *is the set of all nodes on some infinite path of* $\mathcal{T}$*;*
- $d^\alpha(\mathcal{T})$ *is the set of all nodes* x *of* $\mathcal{T}$ *such that for all* $\beta < \alpha$ *there are infinite paths* P, Q *on* $\mathcal{T}$ *with* $x \in P \cap Q$ *and* $P \cup Q \subseteq d^\beta(\mathcal{T})$.

The **Cantor-Bendixson rank** *of* $\mathcal{T}$ *is the first ordinal* α *such that* $d^\alpha(\mathcal{T}) = d^{\alpha+1}(\mathcal{T})$.

Theorem 63 *There is a binary tree* $(T, \preceq, S)$ *such that* (T, S) *has an automatic copy but* $(T, \preceq)$ *does not.*

Proof. Let the tree $(T, \preceq)$ be generated by the following successor structure.

- If x does not contain a 0 then the successors are obtained by appending one digit: $S(1^n, 1^n0)$, $S(1^n, 1^n1)$.
- If x contains a 0 and starts with a 1, that is, if $x = 1v0w$, then the successors of x are obtained by either appending a 0 or rotating to the left: $S(1v0w, 1v0w0)$, $S(1v0w, v0w1)$.
- If x starts with a 0, that is, if $x = 0y$, then x has only the successor obtained by appending a 0: $S(0y, 0y0)$.

The Cantor-Bendixson rank of $\mathcal{T}$ is $\omega + 1$: The n-th derivative $d^n(\mathcal{T})$ is the set containing all strings of the form 1^n and $1^{n+m}0y$ where y contains arbitrary many 0's but at most m 1's. The nodes $1^{n+m}0y$ where y contains exactly m 1's are exactly the nodes on isolated paths of the tree $(d^n(\mathcal{T}), \preceq)$ and therefore omitted from $d^{n+1}(\mathcal{T})$. It follows that $d^\omega = \{1\}^*$ and $d^{\omega+1} = \emptyset$. Thus T has only countably many infinite paths and Cantor-Bendixson rank $\omega + 1$.

Khoussainov, Rubin and Stephan [9] have shown that every automatic tree $(\mathcal{T}, \preceq)$ in the signature of partial orders has finite Cantor-Bendixson rank. Since the Cantor-Bendixson rank does not depend on the presentation, it is impossible to find a presentation of the considered tree where T is regular and the ordering $\preceq$ is automatic. □

7 Conclusion

In this paper we studied the relationship between regularity and definability of relations in automatic structures. On the one hand, we have shown that definable relations do not always capture intrinsic regularity. On the other, we have provided some examples and results in which intrinsic regularity and definability coincide. We have concentrated our study on fragments of arithmetic. Naturally, there are many examples of other fundamental structures, such as binary trees, the group of integers, and the orderings of the rational numbers, in which the interplay between definability and regularity can be studied. It seems that the formal language that we use to define relations (namely first order logic with the added quantifier $\exists^\infty$) is not powerful enough to capture the concept of intrinsic regularity. It would be interesting to investigate questions related to finding formal systems in which definability and intrinsic regularity coincide.

References

1. Christopher J. Ash and Anil Nerode. Intrinsically recursive relations. In *Aspects of effective algebra (Clayton, 1979)*, pages 26–41. Upside Down A Book Co. Yarra Glen, 1981.
2. Veronique Bruyère, Georges Hansel, Christian Michaux and Roger Villemaire. Logic and p-recognizable sets of integers. *Bull. Belg. Math. Soc.*, 1:191–238, 1994.
3. Achim Blumensath. *Automatic Structures.* Diploma thesis, RWTH Aachen, 1999.
4. Samuel Eilenberg, Calvin C. Elgot, and John C. Shepherdson. Sets recognised by n–tape automata. *Journal of Algebra*, 13(4):447–464, 1969.
5. Herbert B. Enderton. *A mathematical introduction to logic.* Academic Press, first edition, 1972.
6. Alexander S. Kechris. *Classical Descriptive Set Theory.* Springer-Verlag, 1995.
7. Bakhadyr Khoussainov and Anil Nerode. Automatic presentations of structures. *Lecture Notes in Computer Science*, 960:367–392, 1995.
8. Bakhadyr Khoussainov and Sasha Rubin. Some thoughts on automatic structures. *Journal of Automata, Languages and Combinatorics.* Accepted.
9. Bakhadyr Khoussainov, Sasha Rubin and Frank Stephan. Automatic partial orders. Manuscript, 2003.