Spectra with Only Unary Function Symbols

Arnaud Durand
arnaud@info.unicaen.fr
GREYC, CNRS UPRESA 6072
Université de Caen
F - 14032 Caen Cedex

Ronald Fagin
fagin@almaden.ibm.com
IBM Almaden Research Center
650 Harry Road
San Jose, California 95120-6099

Bernd Loescher
loescher@lri.fr
Laboratoire de Recherche en Informatique
Université Paris-Sud ORSAY
F - 91405 Orsay Cedex

Appeared in: Annual Conference of the 1997 European Association for Computer Science Logic (CSL'97). In Springer-Verlag Lecture Notes in Computer Science 1414, 1998, ed. M. Nielsen and W. Thomas, pp. 189–202.

Abstract

The spectrum of a first-order sentence is the set of cardinalities of its finite models. This paper is concerned with spectra of sentences over languages that contain only unary function symbols. In particular, it is shown that a set S of natural numbers is the spectrum of a sentence over the language of one unary function symbol precisely if S is an eventually periodic set.

1 Introduction

The *spectrum* of a first-order sentence is the set of cardinalities of its finite models. That is, if φ is a first-order sentence, and if n is a natural number, then n is in the spectrum of φ precisely if there is a structure $\mathcal{A}$ that satisfies φ where the cardinality of the universe of $\mathcal{A}$ is n. The notion of a spectrum was introduced by Scholz [Sc52]. As an example, if φ is a first-order sentence that gives the conjunction of the field axioms (φ says that $+$ and $\times$ are associative and commutative, that $\times$ distributes over $+$, etc.), then it is well-known that the spectrum of φ is the set of powers of primes. In 1952, Scholz [Sc52] posed the problem of characterizing spectra. In 1955, Asser [As55] posed the following key problem, which is sometimes referred to as *Asser's problem*: Is the class of spectra closed under complement? That is, is the complement of every spectrum also a spectrum? These were two of the first problems ever posed in finite model theory. As is discussed in Fagin's survey paper [Fa93], the investigation of Asser's problem is what led Fagin to write his 1973 Ph.D. thesis [Fa73] in finite model theory.

Since the time of Scholz and Asser, there has been only a small amount of progress in our understanding of spectra. Jones and Selman [JS74] (cf. [Fa74]) showed that a set of natural numbers is a spectrum iff it is in NEXPTIME, that is, recognizable by a nondeterministic Turing machine in exponential time. As pointed out in [Fa75, Fa93], it is an open problem as to whether there is any spectrum that is not also a spectrum involving only a single binary relation symbol (that is, the spectrum of a sentence over the language of one binary relation symbol). Fagin [Fa74] proved that there is a spectrum involving only a single binary relation symbol that is NEXPTIME-complete. Therefore, if the class of spectra is not closed under complement, then there is a spectrum involving only a single binary relation symbol whose complement is not a spectrum.

Let us consider the complexity of spectra in restricted languages. It will be convenient in this paragraph to treat natural numbers as being written in unary. It is well-known that spectra involving only unary relation symbols are extremely simple: they are either finite or co-finite sets. As we noted, Fagin [Fa74] showed that there is a spectrum involving only a single binary relation symbol that is NEXPTIME-complete, which means that it is NP-complete when numbers are written in unary. An intermediate case (between spectra involving only unary relation symbols, which are extremely simple, and those involving only a single binary relation symbol, which are extremely complex) are spectra involving only unary function symbols. Are they simple or complex? It follows from results of Durand and Ranaivoson [DR96] that there is a spectrum involving only two unary function symbols that is NP-complete (when numbers are written in unary). This leaves open the complexity of spectra involving only a single unary function symbol. We show that somewhat surprisingly, these spectra are very simple. Specifically, we show that a set S of natural numbers is a spectrum involving only a single unary function symbol if and only if S is eventually periodic, that is, if and only if there are natural numbers N, p with $p > 0$ such that for every n with $n > N$, we have $n \in S$ iff $n + p \in S$. As we shall discuss, it follows that S is a spectrum involving only a single unary function symbol if and only if S is recognizable by a finite automaton (when numbers are written in unary). Furthermore, as we shall also discuss, these results continue to hold even if we allow not only a unary function symbol, but also an arbitrary number of unary relation symbols.

An immediate consequence of our characterization of spectra involving only a single unary function symbol is a resolution of Asser's problem for such sentences: the complement of a spectrum in this restricted language is a spectrum in this restricted language.

Next, we consider the hierarchy based on the number of unary function symbols. The kth level of the hierarchy consists of spectra involving only k unary function symbols. Here, we know only partial results. First, we know that the first two levels of the hierarchy are distinct. This was first shown by Loescher [Lo97], who showed that the set of perfect squares (those numbers of the form n^2, where n is an integer) is in the second level of the hierarchy but not the first level. The fact that the set of perfect squares is not in the first level of the hierarchy follows immediately from our characterization of the first level of the hierarchy as consisting precisely of the eventually periodic sets. Also, the NEXPTIME-complete sets in the second level of the hierarchy (whose existence follows from results of Durand and Ranaivoson [DR96]) are not in the first level, because of our characterization.

We give some evidence that it might be difficult to give a characterization of spectra involving only two unary function symbols. We also give some evidence that the hierarchy based on the number of unary function symbols is strict, and that it might be difficult to prove this. Finally, we show that if S is a spectrum involving only k unary function symbols, then $\{kn \mid n \in S\}$ is a spectrum involving only two unary function symbols.

2 Eventually Periodic Sets

As before, define a set S of natural numbers to be *eventually periodic* if there are natural numbers N, p with $p > 0$ such that for every n with $n > N$, we have $n \in S$ iff $n + p \in S$. The number p is often called the *period*. We now give another characterization of eventually periodic sets, that is a little more convenient for us to work with.

Assume that m, i are natural numbers. Define the *arithmetic series* $A_{m,i}$ to be the set of all numbers of the form $m + \theta i$, for $\theta = 0, 1, 2, \ldots$ In particular, by taking $i = 0$, we see that every singleton set is an arithmetic series. We refer to an arithmetic series $A_{m,i}$ with $i > 0$ as a *nontrivial* arithmetic series. Thus, an arithmetic series is nontrivial precisely if it is not a singleton set. The next proposition is well-known.

Proposition 2.1 *A set of natural numbers is eventually periodic if and only if it is a finite union of arithmetic series.*

Proof: Let S be a finite union of arithmetic series; we now show that S is is eventually periodic. Clearly, S is the union of a finite set and of a set S' that is a finite union $\bigcup_{j \in J} A_{m_j, i_j}$ of nontrivial arithmetic series (so that $i_j > 0$ for each $j \in J$). Define N to be the max of the m_j's, and define p to be the least common multiple of the i_j's. Then for each $n > N$, we have:

$$\begin{aligned} & n \in \bigcup_{j \in J} A_{m_j, i_j} \\ \Leftrightarrow \quad & n \in A_{m_j, i_j} \text{ for some } j \in J \end{aligned}$$

$$\begin{aligned} &\Leftrightarrow \quad n+p \in A_{m_j,i_j} \text{ for some } j \in J \\ &\Leftrightarrow \quad n+p \in \bigcup_{j \in J} A_{m_j,i_j}. \end{aligned}$$

Therefore S', and hence S, is eventually periodic, with period p.

Conversely, let S be an eventually periodic set with parameters N and p. It is clear that S is the union of

- the union of singleton sets of numbers up to N in S, and
- the union of all the arithmetic series $A_{m,p}$ with $N < m \leq N+p$ and $m \in S$.

This union is a finite union of arithmetic series, as desired. □

3 Machinery

The next lemma is a key tool in our characterization of spectra involving only a single unary function symbol.

Lemma 3.1 *Assume that S is a spectrum involving only a single unary function symbol. There are natural numbers k, N such that for each $n \in S$ with $n > N$, there is a natural number i with $0 < i \leq k$ such that $A_{n,i} \subseteq S$.*

This lemma can be proven by using techniques of Loescher [Lo97], although this lemma did not appear in Loescher's paper. In the remainder of this section, we give an informal sketch of the proof of this lemma. Missing details can be obtained from [Lo97].

A *language* $\mathcal{L}$ (sometimes called a *similarity type*, a *signature*, or a *vocabulary*) is a finite set of relation symbols and function symbols, each of which has an arity. An *$\mathcal{L}$-structure* (or *structure over $\mathcal{L}$*, or simply *structure*) is a set A (called the *universe*), along with a mapping associating a relation (resp., function) of the appropriate arity over A (called the *interpretation*) with each relation symbol (resp., function symbol) in $\mathcal{L}$. The structure is called *finite* if A is. In this paper, we restrict our attention to finite structures.

For definitions of a first-order formula and first-order sentence (where, intuitively, the only quantification is over members of the universe, and not over, say, sets of members of the universe) and what it means for a structure $\mathcal{M}$ to *satisfy* a sentence φ, see Enderton [En72] or Shoenfield [Sh67]. If $\mathcal{M}$ satisfies φ, then $\mathcal{M}$ is a *model* of φ. An *$\mathcal{L}$-sentence* is a first-order sentence over the language $\mathcal{L}$. When we refer, for example, to a sentence φ *involving only a single unary function symbol*, we mean that φ is an $\mathcal{L}$-sentence where $\mathcal{L}$ is a language consisting of a single unary function symbol. Similarly, we refer to the spectrum of φ as a *spectrum involving only a single unary function symbol.*

The *quantifier depth* $QD(\varphi)$ of a first-order formula φ is defined recursively as follows: $QD(\varphi) = 0$ if φ is quantifier-free; $QD(\neg\varphi) = QD(\varphi)$; $QD(\varphi_1 \wedge \varphi_2) = \max\{QD(\varphi_1), QD(\varphi_2)\}$; $QD(\exists\varphi) = 1 + QD(\varphi)$.

Two $\mathcal{L}$-structures are said to be *r-equivalent* if they satisfy the same $\mathcal{L}$-sentences of quantifier depth up to r.

For the rest of this section, let $\mathcal{L}$ be the language consisting of a single unary function symbol. We now show how to obtain larger and larger models of an $\mathcal{L}$-sentence, given such a model that is large enough. Assume that φ is an $\mathcal{L}$-sentence of quantifier depth r, that an $\mathcal{L}$-structure $\mathcal{M}$ satisfies φ, and that the universe of $\mathcal{M}$ is "large enough" (that is, bigger than some number depending only on r). The graph of the structure $\mathcal{M}$ (as of any total unary function on a finite set) is a collection of connected components, each of which consists of some mutually disjoint trees whose roots form a cycle (see Figure 1). Intuitively, there is an edge from x to y in the graph if $f(x) = y$, where f is the function in $\mathcal{M}$. By a cutting and pasting argument, we can show that there is an r-equivalent such structure of the same size (that is, whose universe is of the same size) such that

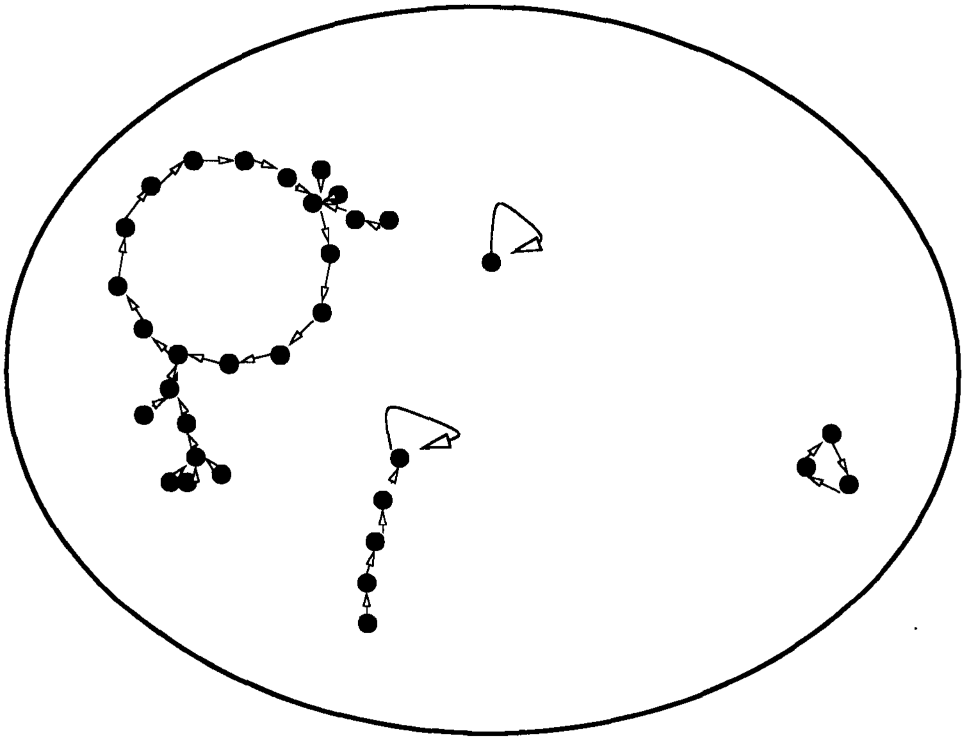

Figure 1: The graph of $\mathcal{M}$

- each tree in this structure is of depth bounded above by a constant, and
- each cycle is of length bounded above by a constant.

("Constant" means here always that the bound depends only on r.) The idea behind this cutting and pasting argument is as follows. First, we deal with shortening the length of long branches. If a branch is sufficiently long, then it has two points a and b that are far away from each other, and have the same "type" (although we do not wish to define Ehrenfeucht-Fraïssé games here, having the same type means intuitively that the points a and b behave identically if either is chosen in a game in the first round and there are $r-1$ rounds left to play). Without loss of generality, assume that a is an ancestor of b. Let a^+ be the son of a that is also an ancestor of b, and let b^+ be the son of b such that (a, a^+) and (b, b^+) have the same type (here, intuitively, this means that the pairs behave identically if either is chosen in a game in the first two rounds and there are $r-2$ rounds left to play). Such a point b^+ is guaranteed to exist by the choice of a and b. As in Figure 2, we remove the edges from b^+ to b and from a^+ to a, and add edges from b^+ to a and from a^+ to b. We thereby shorten the branch, and create a new connected component that contains a long cycle.

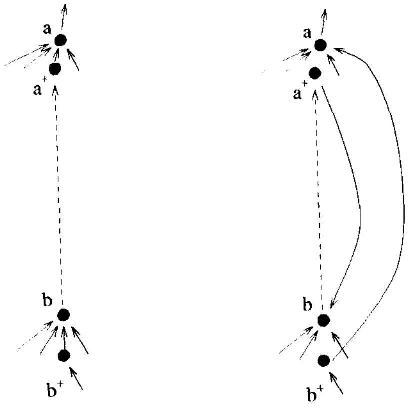

Figure 2: How to shorten branches

After we have shortened all the long branches, we then shorten the long cycles by splitting them, in the same spirit as Fagin, Stockmeyer and Vardi's approach in [FSV95]. If some cycle is sufficiently long, then it has two points a and b that are far away from each other, and have the same type. Let $f(a)$ and $f(b)$ be the images of a and b with respect to the function in $\mathcal{M}$. Then $(a, f(a))$ and $(b, f(b))$ have the same type. As in Figure 3, we remove the edges from b to $f(b)$ and from a to $f(a)$, and add edges from b to $f(a)$ and from a to $f(b)$. (The triangles in Figure 3 represent trees whose roots are on the cycles. Such trees may have roots anywhere on the cycles, including at the points a, b, $f(a)$, and $f(b)$.) We thereby create two shorter cycles.

Then there is a constant c such that:

- If there is a connected component bigger than c, then it must contain a tree of a high width: so high that there must be at least r subtrees that are r-equivalent (when viewed as $\mathcal{L}$-structures), are of size bounded above by a constant s, and whose

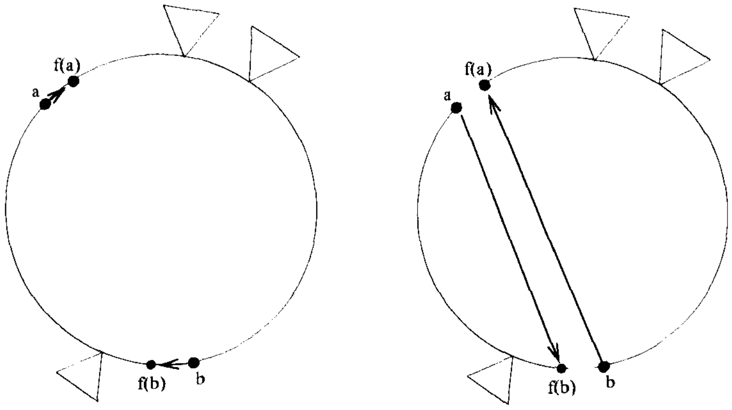

Figure 3: How to split cycles

roots have the same father z (because the number of mutually non-r-equivalent $\mathcal{L}$-structures is bounded above by a constant). The reason for the size bound s is that we select the father z as far as possible from the root of the tree, which guarantees that the width of the subtrees is bounded. Inserting copies of one of these subtrees does not change the r-equivalence class of $\mathcal{M}$. If t is the size of one of these subtrees, then we take i in Lemma 3.1 to be t. We take k in Lemma 3.1 to be bigger than the size bound s.

- If not, then there must be many connected components: enough to have at least r isomorphic ones (because the number of mutually non-isomorphic components of a bounded size is bounded above by a constant). Adding copies of one of these does not change the r-equivalence class of $\mathcal{M}$. If d is the size of one of these connected components, then we take i in Lemma 3.1 to be d. We take k in Lemma 3.1 to be bigger than the constant c that bounds the size of the connected components.

4 Characterization of Spectra Involving only a Single Unary Function Symbol

A set S of natural numbers is said to be *definable in Presburger arithmetic* [En72] if it is definable with a formula involving only addition.

We can now state the main theorem.

Theorem 1 *Let $\mathcal{L}$ be the language that contains only one unary function symbol. Let k be a natural number and let $\mathcal{L}_k$ be the language that contains one unary function symbol and k unary relation symbols.*

Let S be a set of natural numbers. The following properties of S are equivalent:

1. *S is the spectrum of a first-order $\mathcal{L}$-sentence.*

2. *S is the spectrum of a first-order $\mathcal{L}$-sentence, where the interpretation of the function symbol f is restricted to be a permutation.*

3. *S is the spectrum of a first-order $\mathcal{L}_k$-sentence.*

4. *S is eventually periodic.*

5. *S is a finite union of arithmetic series.*

6. *S is definable in Presburger arithmetic.*

7. *When numbers are written in unary, S is recognizable by a finite automaton.*

Proof: (4) ⇔ (7) is well known; see for instance [Ei74, Proposition 1.1, p. 101].

(4) ⇔ (6) is also well known; see for instance [En72, Theorem 32F, p. 188].

(4) ⇔ (5) was shown in Proposition 2.1.

(2) ⇒ (1) is trivial.

(1) ⇒ (3) is also trivial.

(5) ⇒ (2): Let m, i be natural numbers. The arithmetic series $A_{m,i}$ is the spectrum of a sentence that says that except for a set of exactly m points, the remaining points all lie on cycles of size i:

$$\exists x_1 \ldots \exists x_m \forall y[\bigwedge_{1\leq k<l\leq m} (x_k \neq x_l) \wedge ((\bigvee_{1\leq k\leq m} (y = x_i)) \vee (\bigwedge_{1\leq k<i} (f^k(y) \neq y) \wedge (f^i(y) = y)))],$$

where we define $f^k(y)$ recursively by letting $f^k(y)$ be y if $k = 0$, and $f(f^{k-1}(y))$ if $k > 0$. Of course, f can be interpreted by a permutation (which, intuitively, is the identity on the x_i's).

We have shown that each arithmetic series is the spectrum of a first-order $\mathcal{L}$-sentence. A finite union of arithmetic series is then the spectrum of a sentence that is a disjunction of the corresponding sentences.

(1) ⇒ (5): Assume that S is the spectrum of an $\mathcal{L}$-sentence with quantifier depth r. Let k, N be numbers whose existence is guaranteed by Lemma 3.1. Assume $0 < i \leq k$.

Let us say that j is *i-good* if $0 \leq j < i$ and there is $m > N$ with $m \equiv j \bmod i$ and $A_{m,i} \subseteq S$. For each j that is i-good, define $m_{i,j}$ to be the least m where $m \equiv j \bmod i$ and $A_{m,i} \subseteq S$.

We now show that the set of members of S that are greater than N is the union of the sets $A_{m_{i,j},i}$ such that $0 < i \leq k$ and j is i-good. Since this is a finite union of arithmetic series, as are the set of numbers bounded above by N, it follows that S is a finite union of arithmetic series, as desired.

By construction, each such set $A_{m_{i,j},i}$ is a subset of S. So we need only show that if $n > N$ and $n \in S$, then there are i, j such that $0 < i \leq k$ and j is i-good, and $n \in A_{m_{i,j},i}$. Let i be the number whose existence is guaranteed by Lemma 3.1, so that $A_{n,i} \subseteq S$. Let j be such that $0 \leq j < i$ and $n \equiv j \bmod i$. So j is i-good. Since $m_{i,j} \equiv j \bmod i$, it follows by minimality of $m_{i,j}$ that $m_{i,j} \leq n$. Therefore, since $n \equiv m_{i,j} \bmod i$, it follows that $A_{n,i} \subseteq A_{m_{i,j},i}$. Since $n \in A_{n,i}$, we have $n \in A_{m_{i,j},i}$, as desired.

(3) ⇒ (4): Loescher observed [Lo97] that his construction works also in the presence of an arbitrary finite number of unary relation symbols. From this, we obtain an analogue of Lemma 3.1 for the language $\mathcal{L}_k$, and we can repeat the same argument as before. □

We view the equivalence of (1) and (4) as our main result, since it gives a simple characterization of the spectra of sentences involving only a single unary function symbol. Since the set of perfect squares is not eventually periodic, the following corollary is immediate.

Corollary 4.1 [Lo97] *The set of perfect squares is not a spectrum involving only a single unary function symbol.*

Since the set of perfect squares is a spectrum involving only two unary function symbols, this gives us the following result.

Corollary 4.2 [Lo97] *There is a spectrum involving only two unary function symbols that is not a spectrum involving only a single unary function symbol.*

Since the class of eventually periodic sets is closed under complement, we obtain the following additional corollary from the equivalence of (1) and (4) in Theorem 1.

Corollary 4.3 *The class of spectra involving only a single unary function symbol is closed under complement.*

This corollary gives a resolution of Asser's problem when we restrict our attention to spectra involving only a single unary function symbol.

The equivalence of (1) and (2) in Theorem 1 tells us that our characterization continues to hold even if we "tighten up" by demanding that the unary function symbol be interpreted by a permutation. Thus, this equivalence gives a certain sense in which permutations and general unary functions have the same "expressive power". There are other contexts in which this is not the case. For example, in Durand, Lautemann and Schwentick's study of binary NP [DLS96], they showed that there is strictly less expressive power in existentially quantifying over permutations than in existentially quantifying over arbitrary unary functions.

The equivalence of (1) and (3) in Theorem 1 tells us that our characterization continues to hold even if we "loosen up" by allowing, in addition to a unary function symbol, also an arbitrary number of unary relation symbols. Thus, intuitively, adding an arbitrary number of unary relation symbols to the language $\mathcal{L}$ does not increase the expressive power. By contrast, doing the same with the language of addition—which has the same expressive power as $\mathcal{L}$, by the equivalence of (1) and (6) in Theorem 1—does have an effect. In fact, Lynch [Ly82, Lemma 1, p. 134] showed that multiplication is first-order definable in the language of addition and two unary relation symbols.

Since eventual periodicity is such a natural and simple notion, it is not surprising that there are various conditions that arise in mathematics that are characterized by being eventually periodic. Conditions (6) and (7) of Theorem 1 are simply two examples.

5 Allowing Multiple Unary Function Symbols

The equivalence of (1) and (6) in Theorem 1 says that the class of spectra involving only a single unary function symbol coincides with the class of sets definable in Presburger arithmetic. Can we similarly characterize the class of spectra involving only two unary function symbols by means of arithmetic tools? We now give some evidence that such a characterization may be hard to prove.

A set of integers is said to be *rudimentary* (for short, in RUD) if it can be defined with addition, multiplication and variable-bounded quantification. As an example, the following formula (with free variable p) defines the set of prime numbers:

$$\forall x \leq p \, \forall y \leq p \quad [(xy = p) \Rightarrow ((x = 1) \vee (x = p))].$$

It has been proved that every rudimentary set is a spectrum involving only two unary function symbols (see [Ol96]). The converse is still an open problem.

Let us examine the consequences of the converse holding, that is, of the assumption that every spectrum involving only two unary function symbols is a rudimentary set. It can be shown, using results of [DR96] (together with Proposition 5.2 below), that if S is a spectrum, then there exist positive integers h and k such that $hS^k = \{hn^k | n \in S\}$ is in the class of spectra involving only two unary function symbols. Now suppose that latter class corresponds exactly to RUD. Harrow [Ha73] proved that RUD is closed under "polynomial substitution", in which a polynomial is substituted for a variable. Since RUD allows variable-bounded quantification, Harrow's result implies in particular that if hS^k is in RUD, then $S = \{n | (\exists m \leq hn^k)((m = hn^k) \wedge (m \in hS^k))\}$ is also in RUD. Finally, as the complement of every rudimentary set is also a rudimentary set, we obtain that NEXPTIME = co-NEXPTIME. Using directly arguments of Woods [Wo81], it can also be shown that another consequence of the converse holding is that NP $\neq$ co-NP.

Because of the implication that NEXPTIME = co-NEXPTIME, it is probably unlikely (and certainly very hard to prove!) that spectra involving only two unary function symbols are precisely the rudimentary sets. We do not know of any other candidates for a natural number-theoretic characterization of the spectra of sentences involving only two unary function symbols. As we noted, it follows from results of Durand and Ranaivoson [DR96] that there is a spectrum involving only two unary function symbols that is NEXPTIME-complete.

In the remainder of this section, we consider the hierarchy determined by the number of unary function symbols. Corollary 4.2 says that the first two levels of this hierarchy are distinct. We shall present some evidence that the hierarchy is strict, and some evidence that it is hard to prove this.

We first present some evidence that the hierarchy is strict. Grandjean [Gr85, Gr90] considers nondeterministic RAMs that, when given the number n as input, run for cn steps (any of which may be nondeterministic), for some constant c. In Grandjean's model, (a) at each moment every register stores an integer whose value is at most linear in the input n, and (b) the value of a register may be "guessed" in a single nondeterministic step. Grandjean shows that a set S of positive integers is recognizable by such a machine precisely if S is the spectrum of a sentence φ of the form $\forall x \psi$, where ψ is a quantifier-free formula involving only unary function symbols (thus, φ has only a single universal quantifier). If S is a spectrum involving only k unary function symbols (i.e., if only a fixed number k of unary function symbols is needed in formulas to characterize any such set of integers), then it is easy to see that in Grandjean's model, S is recognized by a nondeterministic RAM that on input the positive integer n, runs for at most kn nondeterministic steps and a polynomial in n number of deterministic steps. So we get the following proposition.

Proposition 5.1 *Assume that every spectrum involving only unary function symbols is a spectrum involving only k unary function symbols. Let S be a set of positive integers that is recognized by a nondeterministic RAM that, when given the number n as input, runs for cn steps (nondeterministic and deterministic), for some constant c. Then S is recognized by a nondeterministic RAM that on input n runs for at most kn nondeterministic steps and a polynomial in n number of deterministic steps.*

Intuitively, Proposition 5.1 tells us that a collapse is unlikely, since such a collapse implies that there is some constant k such that cn nondeterministic steps (for arbitrary c) can be simulated by kn nondeterministic steps and a polynomial number of deterministic steps.

We now give some evidence that it will be hard to prove that the hierarchy is strict, and in particular does not collapse to the second level.

Let σ be the language $\{f_1, \ldots, f_m\}$ consisting of m distinct unary function symbols for some m with $m \geq 3$. Let τ be the language $\{f_1, f_2, U\}$ consisting of two distinct unary function symbols and one unary relation symbol. Let (*) be the following statement:

(*) There is a σ-sentence φ whose spectrum does not coincide with the spectrum of any τ-sentence ψ on the set of perfect squares—that is,

$$sp(\varphi) \cap \{\text{squares}\} \neq sp(\psi) \cap \{\text{squares}\}$$

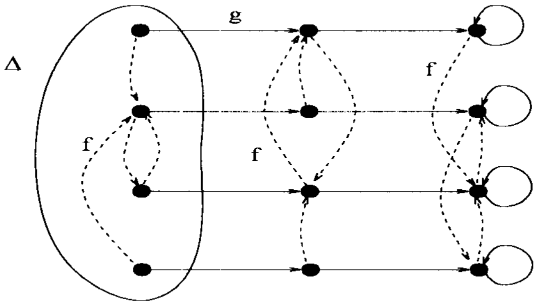

Figure 4: f simulates 3 functions, and g matches corresponding points

Loescher and Sharell [LS] have recently shown that (*) implies that there is a graph property that is in NP but not in binary NP (definable by an existential second order sentence with quantification only over binary relation symbols). So a proof of (*) would resolve the long-standing open problem [Fa75, Fa93] as to whether there is a graph property that is in NP but not in binary NP. Although Loescher and Sharell's result does not directly deal with our hierarchy (because they allow also a unary relation symbol), their result at least provides some evidence that the strictness of our hierarchy might be hard to prove.

We conclude this section with a proposition that gives a weak sense in which k unary function symbols can be "simulated by" only two.

Proposition 5.2 *Let k be a positive integer. Assume that S is a spectrum involving only k unary function symbols. Then $\{kn \mid n \in S\}$ is a spectrum involving only two unary function symbols.*

Proof: Let φ be a first-order sentence involving only the unary function symbols $f_1, \ldots, f_k$, and let f, g be two additional unary function symbols. Let $\Delta(x)$ be the following formula with x as its free variable:

$$\exists z_2 \ldots \exists z_k[(\bigwedge_{i \neq j} z_i \neq z_j) \wedge (g(x) = z_2) \wedge (\bigwedge_{i=2}^{k-1} g(z_i) = z_{i+1}) \wedge (g(z_k) = z_k)].$$

The intuition is that in Figure 4 (which deals with the case $k = 3$, of 3 unary function symbols), $\Delta(x)$ represents the points in the first column, and z_i is the point "corresponding" to x in the ith column.

As before, define $g^i(y)$ recursively to be y if $i = 0$, and $g(g^{i-1}(y))$ if $i > 0$. Let ψ_1 be a sentence that says intuitively that the k sets $g^i(\Delta)$, for $i = 0, \ldots, k-1$, which correspond to the k columns of Figure 4, form a partition of the universe, and that g is one-to-one. Define ψ_2 to be a sentence that says intuitively that $f(g^i(\Delta)) \subseteq g^i(\Delta)$ for $i = 0, \ldots, k-1$. (It is clear how to write such sentences ψ_1 and ψ_2.) Intuitively, the sentence ψ_2 says that f maps points in the ith column into points in the ith column.

Define ψ_3 to be the formula obtained from φ as follows. First, we replace the formula by an equivalent formula where every subformula that is an inequality is of the form $x \neq y$, where x and y are variables (and in particular, do not involve function symbols). We do this by replacing each inequality $t_1 \neq t_2$ by $\exists x \exists y[(t_1 = x) \wedge (t_2 = y) \wedge (x \neq y)]$, where x and y are new variables. Next, we replace the formula by an equivalent formula where no subformula has a function symbol on the right-hand side of an equality. We do this by replacing each subformula $t_1 = f_i(t_2)$ by $\exists x[(t_1 = x) \wedge (f_i(t_2) = x)]$, where x is a new variable. Next, we "de-nest" the f_i's, by recursively replacing each subformula $f_i(f_j(t)) = x$ by $\exists y[(f_j(t) = y) \wedge (f_i(y) = x)]$, where y is a new variable. We have now reduced to the case where every subformula involving a function symbol is of the form $f_i(x) = y$, where x and y are variables. Next, we replace every subformula of the form $f_i(x) = y$ by $f(g^{i-1}(x)) = g^{i-1}(y)$. Intuitively, the function f_i is simulated by the action of f in the ith column of Figure 4. Finally, we relativize every variable x to Δ.

It is not hard to see that if the spectrum of φ is S, then the spectrum of $\psi_1 \wedge \psi_2 \wedge \psi_3$ is $\{kn \mid n \in S\}$. □

This proposition is analogous to Fagin's result [Fa75] that for every spectrum S, there is a positive integer k such that $\{n^k \mid n \in S\}$ is a spectrum involving only a single binary relation symbol. However, Proposition 5.2 is much easier to prove than Fagin's result.

6 Extension to Richer Logics

We can of course consider the spectrum of a sentence σ not only in first-order logic, but in richer logics as well, by again taking the spectrum of σ to be the set of cardinalities of the finite models of σ. Consider statement (**) below, which (in the case when σ is first-order) is simply the equivalence of (1) and (4) in Theorem 1.

(**) A set S of natural numbers is the spectrum of a sentence σ over the language of a single unary function symbol precisely if S is an eventually periodic set.

Let f be a unary function symbol, let $U_1, \ldots, U_k$ be unary relation symbols, and let φ be a first-order sentence over the language $\{f, U_1, \ldots, U_k\}$ (we may write φ as $\varphi(f, U_1, \ldots, U_k)$ to emphasize the language). Then

$$\exists U_1 \ldots \exists U_k \varphi(f, U_1, \ldots, U_k)$$

is a *monadic NP sentence* [FSV95] over the language of a single unary function symbol (since the unary relation symbols are quantified out).

It follows from the equivalence of (3) and (4) in Theorem 1 and from the fact that φ and $\exists U_1 \ldots \exists U_k \varphi$ have the same spectrum that our characterization (**) holds even if σ is allowed to be a monadic NP sentence.

After hearing our results, Gurevich and Shelah [GS] extended our characterization (**) even further, to hold even if σ is allowed to be a monadic second-order sentence

$$Q_1U_1 \dots Q_kU_k\varphi(f, U_1, \dots, U_k),$$

where the Q_i's may be either universal or existential second-order quantifiers.

We now show that our characterization (**) does not hold when we extend yet further by allowing σ to be a sentence $\exists P\varphi(f, P)$, where P is a binary relation symbol. In fact, we shall show that (**) need not hold even when σ is allowed to be a sentence $\exists g\varphi(f, g)$, where g is a unary function symbol. Let $\varphi(f, g)$ be a first-order sentence involving only two unary function symbols whose spectrum is not an eventually periodic set. For example, we can let $\varphi(f, g)$ be the sentence Loescher [Lo97] defined whose spectrum is the set of perfect squares, or we can let $\varphi(f, g)$ be the sentence Durand and Ranaivoson [DR96] showed has an NEXPTIME-complete spectrum. Then the spectrum of $\exists g\varphi(f, g)$ is the same as the spectrum of $\varphi(f, g)$, and hence is not an eventually periodic set.

7 Summary

We show that a set S of natural numbers is the spectrum of a first-order sentence involving only a single unary function symbol precisely if S is an eventually periodic set. We show that is true also if the unary function symbol is restricted to represent a permutation, and this is true also if we allow not only a unary function symbol but also an arbitrary number of unary relation symbols. Finally, we consider the hierarchy that is based on the number of unary function symbols, and obtain some preliminary results.

References

[As55] G. Asser, Das Repräsentantenproblem im Prädikatenkalkül der ersten Stufe mit Identität, *Z. Math. Logik Grundlag. Math.* 1, 1955, pp. 252–263.

[DLS96] A. Durand, C. Lautemann and T. Schwentick, Subclasses of binary NP, *Journal of Logic and Computation*, to appear.

[DR96] A. Durand and S. Ranaivoson, First-order spectra with one binary predicate, *Theoretical Computer Science* 160, 1–2, 1996, pp. 305–320.

[Ei74] S. Eilenberg, *Automata, Languages, and Machines, Vol. A*, Academic Press, New York and London, 1974.

[En72] H. Enderton, *A Mathematical Introduction to Logic*, Academic Press, New York and London, 1972.

[Fa73] R. Fagin, *Contributions to the model theory of finite structures*, Ph.D. Thesis, University of California at Berkeley, 1973.

[Fa74] R. Fagin, Generalized first–order spectra and polynomial–time recognizable sets, in: R. M. Karp, ed., *Complexity of Computation, SIAM-AMS Proc.* 7, 1974, pp. 43–73.

[Fa75] R. Fagin, A spectrum hierarchy, *Z. Math. Logik Grundlag. Math.* 21, 1975, pp. 123–134.

[Fa93] R. Fagin, Finite-model theory—a personal perspective, *Theoretical Computer Science* 116, 1993, pp. 3–31.

[FSV95] R. Fagin, L. Stockmeyer and M. Y. Vardi, On monadic NP vs. monadic co-NP, *Information and Computation* 120, 1, 1995, pp. 78–92.

[Gr85] E. Grandjean, Universal quantifiers and time complexity of random access machines, *Math. Systems Theory* 18, 1985, pp. 171–187.

[Gr90] E. Grandjean, First-order spectra with one variable, *J. Comput. Systems Sci.* 40, 2, 1990, pp. 136–153.

[GS] Y. Gurevich and S. Shelah, *The monadic second-order theory of one unary function*, in preparation.

[Ha73] K. Harrow, *Sub-elementary classes of functions and relations*, Doctoral Dissertation, New York University, Department of Mathematics, 1973.

[JS74] N. G. Jones and A. L. Selman, Turing machines and the spectra of first-order formulas, *J. Symbolic Logic* 39, 1974, pp. 139–150.

[Lo97] B. Loescher, One unary function says less than two in existential second order logic, *Information Processing Letters* 61, 1997, pp. 69–75.

[LS] B. Loescher and A. Sharell, *The expressive power of quantification over functions in existential second order logic*, in preparation.

[Ly82] J. Lynch, Complexity classes and theories of finite models, *Math. Systems Theory* 15, 1982, pp. 127–144.

[O196] F. Olive, *Caractérisation logique des problèmes NP: robustesse et normalisation*, Ph.D. Thesis, Université de Caen, 1996.

[Sc52] H. Scholz, Problem #1: Ein ungelöstes Problem in der symbolischen Logik, *J. Symbolic Logic* 17, 1952, p. 160.

[Sh67] J. R. Shoenfield, *Mathematical Logic*, Addison-Wesley, Reading, MA, 1967.

[Wo81] A. Woods, *Some problems in logic and number theory and their connections*, Ph.D. Thesis, University of Manchester, 1981.