

A note on the variable hierarchy of first-order spectra

Non Peer-reviewed author version

TAN, Tony (2014) A note on the variable hierarchy of first-order spectra.

Handle: <http://hdl.handle.net/1942/16519>

A note on the variable hierarchy of first-order spectra

Eryk Kopczyński*

Tony Tan†

Abstract

The *spectrum* of a first-order logic sentence is the set of natural numbers that are cardinalities of its finite models. Besides its connection with complexity theory, spectra are also closely linked with general combinatorics, as many combinatorial properties are expressible within first-order logic.

In this note we study the hierarchy of first-order spectra based on the number of variables. We show that it forms an infinite hierarchy. However, despite the fact that more variables can express more spectra, we also show that to show whether the first-order spectra are closed under complement, it is sufficient to consider sentences using only three variables and binary relations.

1 Introduction

The spectrum of a first-order sentence ϕ , denoted by $\text{SPEC}(\phi)$, is the set of natural numbers that are cardinalities of finite models of ϕ . Or, more formally, $\text{SPEC}(\phi) = \{n \mid \phi \text{ has a model of size } n\}$. A set is a *spectrum*, if it is the spectrum of a first-order sentence. We let SPEC to denote the class of all spectra.

The notion of the *spectrum* was introduced by Scholz in [7], where he also asked whether there exists a necessary and sufficient condition for a set to be a spectrum. Since its publication, Scholz's question and many of its variants have been investigated by many researchers for the past 60 years. Arguably, one of the main open problems on spectra is the one asked by Asser in [1], known as *Asser's conjecture*, whether the complement of a spectrum is also a spectrum.

Though seemingly unrelated, it turns out that the notion of spectra has a tight connection with complexity theory. In fact, Asser's conjecture is shown to be equivalent to the problem NE vs. CO-NE*, when Jones and Selman [5], as well as Fagin [4] independently showed that a set of integers is a spectrum if and only if its binary representation is in NE. It also immediately implies that if Asser's conjecture is false, i.e., there is a spectrum whose complement is not a spectrum, then $\text{NP} \neq \text{CO-NP}$, hence $\text{P} \neq \text{NP}$. We refer the reader to [3] for a more comprehensive treatment on the spectra problem and its history.

In this paper we study the following hierarchy of spectra. For every integer $k \geq 1$, define

$$\text{SPEC}_k = \{\text{SPEC}(\phi) \mid \phi \text{ uses only up to } k \text{ variables}\}$$

*University of Warsaw, erykk@mimuw.edu.pl

†Hasselt University and Transnational University of Limburg, ptony.tan@gmail.com

*NE is the class of languages accepted by a non-deterministic (and possibly multi tapes) Turing machine with run time $O(2^{kn})$, for some constant $k > 0$.

Obviously we have $\text{SPEC}_1 \subseteq \text{SPEC}_2 \subseteq \dots$.

It was conjectured that the variable hierarchy collapses to three variables, due to the fact that three variables are enough to describe the computation of a Turing machine. For more discussion on these conjectures, see [3].

In this paper we show the opposite: The variable hierarchy has infinite number of levels, that is, for every $k \geq 3$, $\text{SPEC}_k \subsetneq \text{SPEC}_{2k+2}$ (Corollary 3.3).[†] Our proof follows from the following observations.

- To encode a non-deterministic Turing machine with run time $O(2^{kn})$ with first-order logic, $2k + 1$ number of variables are sufficient.
- Checking whether a structure of size N is a model of a k -variable first-order sentence can be done in $O(N^k \log^2(N))$.

Curiously, despite the infinity of the variable hierarchy, by standard padding argument, our proof implies that the spectra are closed under complement if and only if the complement of every spectra of three-variable sentence (using only binary relations) are also spectra (Corollary 2.5). This means that to settle Asser's conjecture, it is sufficient to consider only the three-variable sentences and using only binary relations.

This note is organised as follows. In Section 2 we present a rather loose hierarchy: for every integer $k \geq 3$, $\text{SPEC}_k \subsetneq \text{SPEC}_{4k+3}$. Then in Section 3 we show that by more careful book-keeping, we obtain a tighter hierarchy: for every integer $k \geq 3$, $\text{SPEC}_k \subsetneq \text{SPEC}_{2k+2}$.

2 An easier hierarchy

For a set $A \subseteq \mathbb{N}$ and a language $L \subseteq \{0, 1\}^*$, when we write $A = L$, we mean that L is the set of the binary representation of the numbers in A . Let $\text{NTIME}[2^{kn}]$ to be the class that consists of $L \subseteq \{0, 1\}^*$ such that L is accepted by a nondeterministic (possibly multi tapes) Turing machine (NTM) with run time $O(2^{kn})$. The class NE denotes $\bigcup_{k \geq 0} \text{NTIME}[2^{kn}]$.

For $w \in \{0, 1\}^*$, $|w|$ denotes the length of w and N_w the number whose binary representation is w . Assuming that w does not start with 0, we have $2^{|w|-1} \leq N_w \leq 2^{|w|} - 1$ and hence, $k2^{|w|} \leq 2kN_w$, for every integer $k \geq 0$.

Proposition 2.1 $\text{NTIME}[2^n] \subseteq \text{SPEC}_3$.

Proof. The proof is via standard encoding of an accepting run of an NTM with a grid. For completeness, we present it here.

Let M be an m tape NTM with running time $\leq c2^n$, for some constant c . On input w , the computation of M on w can be pictured as a grid $G_{M,w} = \{1, 2, \dots, 2cN_w\} \times \{1, 2, \dots, 2cN_w\}$, in which every point $(p, q) \in \{1, 2, \dots, 2cN_w\} \times \{1, 2, \dots, 2cN_w\}$ is labeled with either the states of M , the alphabet of M , or the blank symbol $\#$. Moreover, we can assume that the state q can only be used to labeled the points in $\{1, \dots, 2cN_2\} \times \{1\}$. These labels can be viewed as a binary relations on $\{1, 2, \dots, 2cN_w\}$. Let τ be the set of all these binary relations.

By further partitioning the grid $G_{M,w}$ into $4c^2$ number of grids G_1, \dots, G_{4c^2} , where each grid $G_i = \{1, \dots, N_w\} \times \{1, \dots, N_2\}$. Making $4c^2$ copies $\tau_1, \dots, \tau_{4c^2}$ of the τ , we can label each point

[†]Here we should note that it is already known that $\text{SPEC}_1 \subsetneq \text{SPEC}_2 \subsetneq \text{SPEC}_3$, see [3].

(p, q) in G_i with relations from τ_i . Hence, the grid $G_{M,w}$ can be viewed as a grid $\{1, \dots, N_w\}$, where each point is labeled with relations from $\tau_1 \cup \dots \cup \tau_{4c^2}$.

We can further declare a successor and total ordering on $\{1, \dots, N_w\}$ using three variables. It is then straightforward to write a first-order sentence using only three variables that describes that the labelling of each point (x, y) and its neighbours $(x+1, y)$, $(x, y+1)$ and $(x+1, y+1)$ is according to the transitions in M . This completes our proof of Proposition 2.1. ■

Proposition 2.1 can be generalised to $\text{NTIME}[2^{kn}]$ as stated in the following theorem.

Theorem 2.2 *For every integer $k \geq 1$, $\text{NTIME}[2^{kn}] \subseteq \text{SPEC}_{2k+1}$.*

Proof. The proof is a straightforward extension of the one in Proposition 2.1. Let M be an NTM running in time $\leq c2^{kn}$, for some integer $c, k \geq 1$. On every $w \in L(M)$, the computation of M on w can be represented as a $(c2^k N_w^k \times c2^k N_w^k)$ -square G , which in turn, can be partitioned into $c^2 2^{2k}$ number of $(N_w^k \times N_w^k)$ -squares.

Numbers in $\{1, \dots, N_w^k\}$ can be represented as vectors $(p_1, \dots, p_k) \in \{1, \dots, N_w\}^k$. The successor relation $\text{suc}(x_1, \dots, x_k, y_1, \dots, y_k)$ on $\{1, \dots, N_w\}^k$ can be described as $x_1 = y_1, \dots, x_{i-1} = y_{i-1}, y_i = x_i + 1$ and $y_{i+1} = \dots = y_k = 1$, where i is the smallest index such that $x_i \neq N_w$, and it requires one more new variable. Altogether it uses $2k + 1$ variables. This completes our proof of Theorem 2.2. ■

Theorem 2.3 *For every integer $k \geq 1$, $\text{SPEC}_k \subseteq \text{NTIME}[n2^{2kn}]$.*

Proof. Let φ be an FO sentence using only k variables. Let m be the number of subformulas in φ . We are going to construct an m -tape NTM M that accepts $\text{SPEC}(\varphi)$ with running time $\leq c2^{2kn}$, for some constant $c \geq 1$.

It works as follows. Since φ uses only k variables, we may assume that it uses only relations of arity $\leq k$. Let $w \in \{0, 1\}^*$ be the input string and n be its length. It does the following.

- It computes the number $N = N_w$.
- For each relation R of arity $l \leq k$, it guesses all \bar{a} where $R(\bar{a})$ holds and keep them in a separate tapes for each relation R . This step takes $O(n \cdot N^l) = O(n \cdot 2^{ln})$ steps.
- By induction on the subformula $\psi(z_1, \dots, z_l)$ of φ , where $l \leq k$, it guesses all $\bar{a} \in \{1, \dots, N\}^l$ such that $\psi(\bar{a})$ holds. It simply guesses \bar{a} and check whether $\psi(\bar{a})$ holds. This steps takes $O(N^l \cdot n \cdot N^l) = O(nN^{2l}) = O(n2^{2kn})$.
- Finally, M accepts w if and only if φ holds.

Since φ is fixed, the number m of the subformulae is fixed. Hence, the machine M runs in $O(n2^{2kn})$ nondeterministic time (on m tapes). ■

Corollary 2.4 *For every integer $k \geq 3$, $\text{SPEC}_k \subsetneq \text{SPEC}_{4k+3}$.*

Proof. The strict inclusion follows from

$$\text{SPEC}_k \subseteq \text{NTIME}[n2^{2kn}] \subsetneq \text{NTIME}[2^{(2k+1)n}] \subseteq \text{SPEC}_{2(2k+1)+1} = \text{SPEC}_{4k+3}.$$

The first inclusion follows from Theorem 3.2 and the third from Theorem 2.2. The second strict inclusion follows from nondeterministic time hierarchy theorem [2, Theorem 3.2]. ■

The following corollary shows that to settle Asser's conjecture, it is sufficient to consider sentences using three variables and binary relations.

Let $\text{CO-SPEC}_3^{\text{bin}} = \{\mathbb{N} - S \mid S = \text{SPEC}(\phi) \text{ and } \phi \text{ uses three variables and binary relations}\}$.

Corollary 2.5 $\text{NE} = \text{CO-NE}$ if and only if $\text{CO-SPEC}_3^{\text{bin}} \subseteq \text{SPEC}$.

Proof. The “only if” direction is trivial. The “if” direction is as follows. Suppose $\text{CO-SPEC}_3^{\text{bin}} \subseteq \text{SPEC}$. Since $\text{NTIME}[2^n] \subseteq \text{SPEC}_3$ (and uses only binary relations), this means that for every $A \in \text{NTIME}[2^n]$, the complement $\mathbb{N} - A \in \text{SPEC}$, and hence, also $\mathbb{N} - A \in \text{NE}$. By padding argument, this implies that for every set $A \in \text{NE}$, the complement $\mathbb{N} - A$ also belongs to NE . ■

3 A finer hierarchy

In this section we are going to present a finer hierarchy of the spectra: For every integer $k \geq 3$, $\text{SPEC}_k \subsetneq \text{SPEC}_{2k+2}$. The outline of the proof follows the one in the previous subsection.

Theorem 3.1 For every integer $k \geq 3$, $\text{NTIME}[2^{(k+\frac{1}{2})n}] \subseteq \text{SPEC}_{2k+1}$.

Proof. The idea is as follows. Let M be an NTM with run time $O(2^{(k+\frac{1}{2})n})$. Its computation on an input w can be represented as a grid $G_{M,w} = \left(\{1, \dots, N_w\}^k \times \{1, \dots, \sqrt{N_w}\} \right) \times \left(\{1, \dots, N_w\}^k \times \{1, \dots, \sqrt{N_w}\} \right)$.

Using one variable, we can represent the product $\{1, \dots, \sqrt{N_w}\} \times \{1, \dots, \sqrt{N_w}\}$. Each number $1 \leq x \leq N_w$ represents the point (i, j) , where $x = (i-1)\sqrt{N_w} + j$ and $i, j \in \{1, \dots, \sqrt{N_w}\}$.

The neighbourhood of a point $(\bar{p}, \bar{q}, r) = (p_1, \dots, p_k, q_1, \dots, q_k, r)$ representing a point in the grid $G_{M,w}$ are defined as follows.

- If $r + \sqrt{N_w} \leq N_w$, then the neighbour are the points $(\bar{p}, \bar{q}, r + \sqrt{N_w})$, $(\bar{p}, \bar{q}, r + 1)$ and $(\bar{p}, \bar{q}, r + \sqrt{N_w} + 1)$.
- If $r + \sqrt{N_w} > N_w$ and $r + 1 \leq N_w$, then the neighbour are $(\bar{p}', \bar{q}, r + \sqrt{N_w} - N_w)$, $(\bar{p}, \bar{q}, r + 1)$ and $(\bar{p}', \bar{q}, r + \sqrt{N_w} - N_w + 1)$, where \bar{p}' is the successor of \bar{p} , as defined in the proof of Theorem 2.2.
- If $r + 1 > N_w$, then the neighbour are $(\bar{p}', \bar{q}, \sqrt{N_w})$, $(\bar{p}, \bar{q}', (\sqrt{N_w} - 1)\sqrt{N_w} + 1)$ and $(\bar{p}', \bar{q}', 1)$, where \bar{p}' and \bar{q}' are the successors of \bar{p} and \bar{q} , respectively, as defined in the proof of Theorem 2.2.

To define such neighbourhood, we need to define the following relations.

- The relation $\text{ADD}(x, y, z)$, which holds if and only if $x + y = z$.
- The relation $\text{SQUARE}(x, y)$, which holds if and only if $y = x^2$.
- The relation $\text{DOUBLE}(x, y)$, which holds if and only if $y = 2x$.
- The relation $\text{ADD-SQ-MAX}(x, y)$, which holds if and only if
 - $y = x + \sqrt{N}$, if $x + \sqrt{N} \leq N$,
 - $y = x + \sqrt{N} - N$, if $x + \sqrt{N} < N$,

where N is the maximal element.

All these relations can be defined in first-order sentences using ≤ 8 variables, which is $\leq 2(k + \frac{1}{2}) + 1$, for each integer $k \geq 3$. This completes our proof. \blacksquare

Now, we will show how to compute k -variable spectra effectively.

Theorem 3.2 *For every integer $k \geq 1$, $\text{SPEC}_k \subseteq \text{NTIME}[n^2 2^{kn}]$.*

Proof. Let φ be an FO sentence using only k variables. We are going to construct a multi-tape NTM M that accepts $\text{SPEC}(\varphi)$ with running time $O(n^2 2^{kn})$.

We need the following terminology. Let N be a positive integer. Let T be a d -**dimensional boolean array** of size N . For $1 \leq n_d, \dots, n_1 \leq N$, we denote by $T[n_d, \dots, n_1]$, which is either 0 or 1, the content of the entry (n_d, \dots, n_1) in T . The array is written in a Turing machine tape in the following form: $1, \dots, 1 : T[1, \dots, 1] \$ \dots \$ N, \dots, N : T[N, \dots, N]$, where the number $1, \dots, N$ are written in binary form. The length of the the array T is $O(N^d d \log(N))$.

Claim 1 *Let T be a 2 dimensional boolean array of size N . It is possible to construct the transpose of T in time $O(N^2 n^2)$ on a multi-tape Turing machine, where $n = \log(N)$. That is, if a tape T contains a two-dimensional array, it is possible to write a two-dimensional array on another tape T' , such that $T'[x, y] = T[y, x]$.*

Proof. Intuitively, we use the well known radix sort algorithm to sort the contents of the tape T by their second coordinate. It is well known how to implement a binary counter on a Turing machine so that each step takes amortized $O(1)$ time. The full algorithm is as follows.

We loop x from 1 to N , and y from 1 to N . In each iteration, we read the entry $T[x, y]$. We copy it to a new tape S , followed by the (big endian) binary representation of y in n bits.

The tape S now contains N^2 entries, each of them containing $n + 1$ bits. For $j = 1$ to n , we sort these records by their $n + 2 - j$ -th symbol in a stable way (which is done by first taking all records where this symbol is 0 and copying them to a new tape, then doing the same for all records where this symbol is 1, and then replacing the contents of the old tape by the contents of the new tape). After this operation, our values of $T[x, y]$ will be sorted by y , and values with the same y will keep the original order given by x . Since we have sorted the tape S of length $N^2(n + 1)$ n times, this step takes time $O(N^2 n^2)$. Now, we copy the tape S to T' , removing the binary representations of values of y . As required, $T[x, y]$ has been moved to $T'[y, x]$. \blacksquare

Claim 2 Let σ be a fixed permutation on $\{1, \dots, d\}$. Let T be a d dimensional boolean array of size N . It is possible to construct the transpose of T according to σ in time $O(N^d n^2)$ on a multi-tape Turing machine, where $n = \log(N)$. In other words, it is possible to write a d dimensional array T' on another tape such that $T'[x_1, x_2, \dots, x_d] = T[x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(d)}]$.

Proof. Each permutation σ can be decomposed into at most d number of transpositions $(a_d b_d) \cdots (a_1 b_1)$. We perform each permutation $(a_i b_i)$ starting from $i = 1$ according to the claim above, which takes $O(N^d d n^2)$. Since we repeat the process d times, altogether it takes $O(N^d d^2 n^2) = O(N^d n^2)$. ■

We will show how to use this to effectively check whether $N \in \text{SPEC}(\varphi)$. Let m be the number of subformulae in φ . Since there are only at most k variables, we can assume that every relation in φ has arity at most k .

First, we guess \mathcal{A} , a structure of size N . We can assume that the universe is $\{1, \dots, N\}$. We designate a tape T_R for each relation R which appears in the formula; this tape will contain a d -dimensional array of bits, where d is the arity of R . $T_R[x_1, x_2, \dots, x_d]$ will be 1 if $\mathcal{A} \models R(x_1, \dots, x_d)$, and 0 otherwise.

Now, for each subformula ϕ of φ in bottom-up order, on the tape T_ϕ we will calculate the tuples for which ϕ is true, in similar way: if ϕ has d free variables, $T_\phi[x_1, x_2, \dots, x_d]$ will be 1 if $\mathcal{A} \models \phi(x_1, \dots, x_d)$, and 0 otherwise. It is straightforward to calculate the contents of the tape T_ϕ in time $O(N^k \log N)$: for atomic formulae, just copy T_R ; for $\phi = \neg\phi'$, just negate the contents of $T_{\phi'}$; for $\forall x_1 \phi'(x_1, \dots, x_d)$, just replace a sequence of n ones in $T_{\phi'}$ with 1 in T_ϕ , and all other sequences of n bits with 0; and so on. In some cases we need to adjust the order in variables, but Claim 2 guarantees that we can do it quickly.

Finally, we calculate the tape T_φ for the formula φ itself, as a 0-dimensional array, that is, a single bit. We accept if this tape contains the bit 1.

Since m , the number of subformulae in φ , is fixed, the whole algorithm runs in time $O(N^k n^2) = O(2^{n^k} n^2)$. The constant factor and the number of tapes depend on m . ■

Corollary 3.3 For every integer $k \geq 3$, $\text{SPEC}_k \subsetneq \text{SPEC}_{2k+2}$.

Proof. The strict inclusion follows from

$$\text{SPEC}_k \subseteq \text{NTIME}[n2^{kn}] \subsetneq \text{NTIME}[2^{(k+\frac{1}{2})n}] \subseteq \text{SPEC}_{2(k+\frac{1}{2})+1} = \text{SPEC}_{2k+2}.$$

The first inclusion follows from Theorem 3.2 and the third from Theorem 3.1. The second strict inclusion follows from nondeterministic time hierarchy theorem [2, Theorem 3.2]. ■

References

- [1] G. Asser. *Das Repräsentantenproblem in Prädikatenkalkül der ersten Stufe mit Identität.* Zeitschrift für mathematische Logik und Grundlagen der Mathematik, 1: 252–263 (1955).
- [2] S. Arora, B. Barak. *Computational Complexity – A Modern Approach.* Cambridge University Press, 2009.

- [3] A. Durand, N. D. Jones, J. A. Makowsky, M. More. *Fifty years of the spectrum problem: survey and new results*. Bulletin of Symbolic Logic 18(4): 505–553 (2012).
- [4] R. Fagin. *Contributions to the model theory of finite structures*. Ph.D. thesis, University of California, Berkeley, 1973.
- [5] N. Jones, A. Selman. *Turing machines and the spectra of first-order formulas*. The Journal of Symbolic Logic, 39: 139–150 (1974). (Also in STOC 1972.)
- [6] J. Lynch. *Complexity classes and theories of finite models*. Mathematical Systems Theory 15(2): 127–144 (1982).
- [7] H. Scholz. *Ein ungelöstes Problem in der symbolischen Logik*. The journal of Symbolic Logic, 17: 160 (1952).