On the Hierarchy Classes of Finite Ultrametric Automata

Rihards Krišlauks and Kaspars Balodis

University of Latvia Faculty of Computing

January 27, 2015

Ultrametric finite automata and ultrametric Turing machines

- Introduced by Freivalds in 2012.
- "[..] using *p*-adic numbers is not merely one of many possibilities to generalize the definition of deterministic algorithms but rather the only remaining possibility not yet explored."

Probabilities

- Pascal and Fermat believed that every event of indeterminism can be described by a real number between 0 and 1 called probability.
- Quantum physics introduced a description in terms of complex numbers called amplitude of probabilities and later in terms of probabilistic combinations of amplitudes most conveniently described by density matrices.
- String theory, chemistry and molecular biology have introduced p-adic numbers to describe measures of indeterminism.

p-adic numbers

- For any given prime p the field $\mathbb{Q}_p$ of p-adic numbers is a completion of rational numbers.
- p-adic numbers cannot be linearly ordered.
- In 1916 Alexander Ostrowski proved that any non-trivial absolute value on the rational numbers $\mathbb{Q}$ is equivalent to either the usual real absolute value or a p-adic absolute value.
- Helmut Hasse's local-global principle states that certain types of equations have a rational solution if and only if they have a solution in the real numbers and in the p-adic numbers for each prime p.

Motivation for our paper

- Balodis et al. (2013) showed that regularized ultrametric automata recognize exactly the set of regular languages.
- Holzer (2009), Yao (1978), Monien (1980), Macarie (1995) show results for deterministic, nondeterministic and probabilistic multihead finite automata in the two-way and one-way cases.

Definitions – p-norm

For every non-zero rational number α there exists a unique prime factorization $\alpha = \pm 2^{\alpha_2} 3^{\alpha_3} 5^{\alpha_5} 7^{\alpha_7} \cdots$ where $\alpha_i \in \mathbb{Z}$.
The p-adic absolute value (also called the p-**norm**) of a rational number $\alpha = \pm 2^{\alpha_2} 3^{\alpha_3} 5^{\alpha_5} 7^{\alpha_7} \cdots$ is

$$\|\alpha\|_p = \begin{cases} p^{-\alpha_p}, & \text{if } \alpha \neq 0 \\ 0, & \text{if } \alpha = 0. \end{cases}$$

Definitions – ultrametric automata

A finite one-way p-ultrametric one-head automaton ($1u_pfa$ or $1u_pfa(1)$) is a sextuple $\langle S, \Sigma, s_0, \delta, Q_A, Q_R \rangle$ where

- S is a finite set—the set of states,
- Σ is a finite set ($\$ \notin \Sigma$)—input alphabet,
- $s_0 : S \to \mathbb{Q}_p$ is the initial amplitude distribution,
- $\delta : (\Sigma \cup \{\$\}) \times S \times S \to \mathbb{Q}_p$ is the transition function,
- $Q_A, Q_R \subseteq S$ are the sets of accepting and rejecting states, respectively.

The amplitude distribution after processing the i-th symbol is denoted as s_i, with $s_i(y) = \sum_{x \in S} s_{i-1}(x) \cdot \delta\left(w_i, x, y\right)$ for every $y \in S$.
If $\sum_{x \in Q_A} \left\| s_{n+1}(x) \right\|_p > \sum_{x \in Q_R} \left\| s_{n+1}(x) \right\|_p$, then the word w is said to be accepted, otherwise—rejected.

Results – $1u_pfa(1)$ vs $1nfa(k)$

Let $n = \binom{k}{2} + 1$.

$$L_k = \{w_1 1 w_2 1 \ldots 1 w_{2n} | w_i \in \{0^m | m \geq 1\} \wedge w_i = w_{2n-i+1}\}.$$

Theorem

1. *For every prime p there exists a $1u_pfa(1)$ that recognizes L_k,*
2. *L_k cannot be recognized by any $1nfa(k)$.*

Proof – used constructions

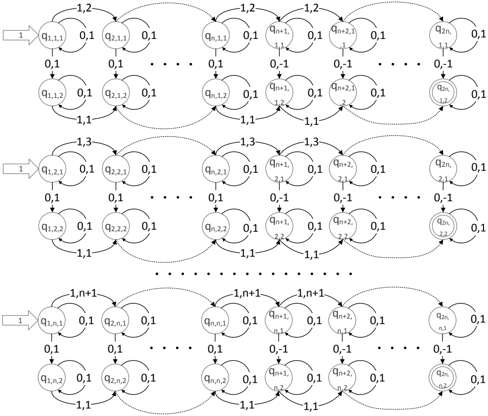

Figure : Automaton for recognizing $0^n10^m10^h1\cdots10^h10^m10^n$. Double-circled states are rejecting. Large arrows with labels in them show the amplitude distribution when the automaton starts. Small labelled arrows show transitions. A label (a, b) indicates that if the automaton reads a, transition with amplitude b should be made.

Proof – used constructions

$$\begin{cases} a_1 + a_2 \cdot 2 + a_3 \cdot 2^2 + \cdots + a_n \cdot 2^{n-1} - a_{n+1} \cdot 2^{n-1} - a_{n+2} \cdot 2^{n-2} - \cdots - a_{2n} = 0 \\ a_1 + a_2 \cdot 3 + a_3 \cdot 3^2 + \cdots + a_n \cdot 3^{n-1} - a_{n+1} \cdot 3^{n-1} - a_{n+2} \cdot 3^{n-2} - \cdots - a_{2n} = 0 \\ \cdots \\ a_1 + a_2 \cdot (n+1) + a_3 \cdot (n+1)^2 + \cdots + a_n \cdot (n+1)^{n-1} - a_{n+1} \cdot (n+1)^{n-1} \\ \qquad - a_{n+2} \cdot (n+1)^{n-2} - \cdots - a_{2n} = 0 \end{cases}$$

rewriting

$$\begin{cases} (a_1 - a_{2n}) + 2 \cdot (a_2 - a_{2n-1}) + 2^2 \cdot (a_3 - a_{2n-2}) + \cdots + 2^{n-1} \cdot (a_n - a_{n+1}) = 0 \\ (a_1 - a_{2n}) + 3 \cdot (a_2 - a_{2n-1}) + 3^2 \cdot (a_3 - a_{2n-2}) + \cdots + 3^{n-1} \cdot (a_n - a_{n+1}) = 0 \\ \cdots \\ (a_1 - a_{2n}) + (n+1) \cdot (a_2 - a_{2n-1}) + (n+1)^2 \cdot (a_3 - a_{2n-2}) + \cdots \\ \qquad + (n+1)^{n-1} \cdot (a_n - a_{n+1}) = 0 \end{cases}$$

Two-way multi-head automata

- Monien (1980) and Macarie (1995) show that deterministic and probabilistic two-way finite automata with k heads are weaker than those with $k + 1$ heads.
- We show that the same results hold for Ultrametric multi-head finite automata.

Two-way multi-head ultrametric automata – separation

By $\widehat{C}$, we denote the subset of a language class C containing only the words in the form $1^{2^n}, n \in \mathbb{N}$, more precisely $\widehat{C} = \left\{L \in C | \forall x \in L \; \exists n \in \mathbb{N} : x = 1^{2^n}\right\}$

Theorem

For every natural number k and prime p:

$$2\widehat{U_p FA(k)} \subsetneq \widehat{U_p TM}.$$

We construct a special p-ultrametric Turing machine with 2 tapes and *log*-space space complexity called $\mathcal{T}$. We show that its recognized language cannot be recognized by a p-ultrametric automata with k heads for any k.

Two-way multi-head ultrametric automata – simulation

We use the function
$f_k : \left\{1^{2^n} | n \in \mathbb{N}\right\} \to \left\{1^{2^n} | n \in \mathbb{N}\right\}, \text{ where } f_k(1^{2^n}) = 1^{2^{k \cdot n}}$. Which is the same as one used by Macarie (1995) and Monien (1980).
When f_k is applied to a language, we refer to the following function: $f_k(L) = \{f_k(x) | x \in L\}$.

Lemma

For every language $L \in \widehat{U_pTM}$ that is recognized by a 2-tape u_ptm in logarithmic space, there exists a natural number u such that: $f_u(L) \in 2\widehat{U_pFA}(3)$.

We show how a u_ptm denoted by $\mathcal{T}$ that recognizes L can be transformed into a u_ptm called $\mathcal{T}'$, which can then be replaced by a p-ultrametric 3 register machine. From this, it easily follows that there exists a $2u_pfa(3)$ that recognizes a "stretched" variant of L, where stretching is done by f_u.

Two-way multi-head ultrametric automata

Lemma

For all languages $L \in \widehat{U_pTM}$ *and all* $u, v \geq 1, u, v \in \mathbb{N}$:
$f_u(L) \in 2\widehat{U_pFA}(v) \Rightarrow L \in 2\widehat{U_pFA}(u \cdot v)$.

Two-way multi-head ultrametric automata

Lemma

For every language $L \in \widehat{U_p TM}$ *and every* $u > v > 1, u, v \in \mathbb{N}$*:*

$$f_{u+1}(L) \in 2\widehat{U_p FA(v)} \Rightarrow f_u(L) \in 2\widehat{U_p FA(v+1)}.$$

Two-way multi-head ultrametric automata – hierarchy classes

Theorem

For all $k \geq 2 \in \mathbb{N}$:

$$2\widehat{U_pFA(k)} \subsetneqq 2\widehat{U_pFA(k+1)}.$$

We prove from the contrary by showing that if there exists such $h \geq 2$ that $2\widehat{U_pFA(h)} = 2\widehat{U_pFA(h+1)}$, it implies $2\widehat{U_pFA(h \cdot (h+1))} = \widehat{U_pTM}$, which contradicts $2\widehat{U_pFA(k)} \subsetneqq \widehat{U_pTM}$.

Questions?