A Note on Two-dimensional Probabilistic Turing Machines

Naoki Tanimoto (University of Tokyo, Department of Computer Science, and Graduate School of Information Science, University of Tokyo)

1. Introduction

The classes of sets recognized by (one-dimensional) probabilistic finite automata and probabilistic Turing machines have been studied extensively [3-6,12-14,18,23]. As far as we know, however, there is only one literature concerned with probabilistic automata on a two-dimensional tape [17]. In [17], we introduced two-dimensional probabilistic finite automata (2-pfa's), and showed that

(i) the class of sets recognized by 2-pfa's with error probability less than \( \frac{1}{\log n} \), 2-PFA, is incomparable with the class of sets accepted by two-dimensional alternating finite automata (2-afa's) [9], and

(ii) 2-PFA is not closed under row concatenation, column concatenation, row + and column + operations in [21].

We believe that it is quite promising to investigate probabilistic machines on a two-dimensional tape.

The classes of sets accepted by two-dimensional (deterministic, nondeterministic, and alternating) finite automata and Turing machines have been studied extensively [1,8-11,15,16,19,22]. In this paper, we introduce a two-dimensional probabilistic Turing machine (2-ptm), and investigate several properties of the class of sets of square tapes recognized by 2-ptm's with error probability less than \( \frac{1}{\log n} \) and with sublogarithmic space.

Section 2 gives some definitions and notations necessary for this paper.

Let 2-PTM\({}^2(L(n))\) be the class of sets of square tapes recognized by \( L(n) \) space-bounded 2-ptm's with error probability less than \( \frac{1}{\log n} \). (See Section 2 for the definition of \( L(n) \) space-bounded 2-ptm's.)

In Section 3, we investigate a relationship between 2-afa's and 2-ptm's with sublogarithmic space, and show that there is a set in 2-AFA\(^s\), but not in 2-PTM\(^2(L(n))\) with \( L(n) = o(\log n) \), where 2-AFA\(^s\) denotes the class of sets of square tapes accepted by 2-afa's. As a corollary of this result, it follows that there is a set in 2-AFA\(^s\), but not recognized by any 2-pfa with error probability less than \( \frac{1}{\log n} \). The partially solves an open problem in [17]. Unfortunately, it is still unknown whether there is a set of square tapes recognized by a 2-pfa with error probability less than \( \frac{1}{\log n} \), but not in 2-AFA\(^s\).

In Section 4, we investigate a space hierarchy of 2-ptm's with error probability less than \( \frac{1}{\log n} \) and with sublogarithmic space. We give several results and introduce a new class of 2-ptm's that solve a problem in [17].

Section 5 shows that if \( L(n) \) is space-constructible by a two-dimensional Turing machine, \( \log \log n < L(n) \leq \log n \) and \( L'(n) = o(L(n)) \), then there is a set of square tapes accepted by a strongly sublogarithmic space-bounded two-dimensional deterministic Turing machine, but not in 2-PTM\(^2(L'(n))\). As a corollary of this result, it follows that 2-PTM\(^2((\log \log n)^k) \supsetneq 2-PTM\(^2((\log \log n)^{k+1}) \) for any positive integer \( k \geq 1 \).

2. Preliminaries

Let \( \Sigma \) be a finite set of symbols. A two-dimensional tape over \( \Sigma \) is a two-dimensional rectangular array of elements of \( \Sigma \). The set of all the two-dimensional tapes over \( \Sigma \) is denoted by \( \Sigma^{(2)} \). Given a tape \( x \) in \( \Sigma^{(2)} \), we let \( l_1(x) \) be the number of rows and \( l_2(x) \) be the number of columns. For each \( m, n \geq 1 \), let \( \Sigma^{m \times n} = \{ x \in \Sigma^{(2)} \mid l_1(x) = m \land l_2(x) = n \} \). If \( 1 \leq i_1 \leq l_1(x) \) for \( k = 1, 2 \), we let \( x(i_1, i_2) \) denote the symbol in \( x \) with coordinates \( (i_1, i_2) \). Furthermore, we define \( x(i_1, i_2) = 0, 1 \) only when \( 1 \leq i_1 \leq i_1' \leq l_1(x) \) and \( 1 \leq i_2 \leq i_2' \leq l_2(x) \), as the two-dimensional tape \( z \) satisfying the following (i) and (ii):

(i) \( l_1(z) = l_1'(i_1) = i_1'' - i_1 + 1 \) and \( l_2(z) = l_2'(i_2) = i_2'' - i_2 + 1 \);  
(ii) for each \( i, j (1 \leq i \leq l_1(z), 1 \leq j \leq l_2(z)) \), \( z(i, j) = x(i_1 + i - 1, i_2 + j - 1) \).

We next introduce a two-dimensional probabilistic Turing machine which is a natural extension of a two-way probabilistic Turing machine [3, 4] to two dimension. Let \( S \) be a finite set. A coin-tossing distribution on \( S \) is a mapping \( \psi \) from \( S \) to \( \{0, \frac{1}{2}, 1\} \) such that \( \Sigma_{a \in S} \psi(a) = 1 \). The mapping means "choose a with probability \( \psi(a) \)."
A two-dimensional probabilistic Turing machine (denoted by 2-ptm) is a 7-tuple $M = (Q, \Sigma, \Gamma, \delta, q_0, q_a, q_r)$, where $Q$ is a finite set of states, $\Sigma$ is a finite input alphabet ($\# \notin \Sigma$ is the boundary symbol), $\Gamma$ is a finite storage tape alphabet ($B \in \Gamma$ is the blank symbol), $\delta$ is a transition function, $q_0 \in Q$ is the initial state, $q_a \in Q$ is the accepting state, and $q_r \in Q$ is the rejecting state. As shown in Fig.1, the machine $M$ has a read-only rectangular input tape over $\Sigma$ surrounded by the boundary symbols # and has one semi-infinite storage tape, initially blank. The transition function $\delta$ is defined on $(Q \times \{q_a, q_r\}) \times (\Sigma \cup \{\#\}) \times \Gamma$ such that for each $q \in Q - \{q_a, q_r\}$, each $\sigma \in \Sigma \cup \{\#\}$ and each $\gamma \in \Gamma$, $\delta(q, \sigma, \gamma)$ is a coin-tossing distribution on $Q \times (\Gamma - \{\delta\}) \times \{\text{Left, Right, Up, Down, Stay}\} \times \{\text{Left, Right, Stay}\}$, where $\text{Left}$ means "moving left", Right "moving right", Up "moving up", Down "moving down" and Stay "staying there". The meaning of $\delta$ is that if $M$ is in state $q$ with the input head scanning the symbol $\sigma$ and the storage tape head scanning the symbol $\gamma$, then with probability $\delta(q, \sigma, \gamma)$, $\delta(q', \gamma', d_1, d_2)$ the machine enters state $q'$, rewrites the symbol $\gamma'$ by the symbol $\gamma$, either moves the input head one symbol in direction $d_1$ if $d_1 \in \{\text{Left, Right, Up, Down}\}$ or does not move the input head if $d_1 = \text{Stay}$, and either moves the storage tape head one symbol in direction $d_2$ if $d_2 \in \{\text{Left, Right}\}$ or does not move the storage tape head if $d_2 = \text{Stay}$.

Given an input tape $x \in \Sigma^2$, $M$ starts in the initial state $q_0$ with the input head on the upper left-hand corner of $x$, with all the cells of the storage tape blank and with the storage tape head on the left end of the storage tape. The computation of $M$ on $x$ is then governed (probabilistically) by the transition function $\delta$ until $M$ either accepts by entering the accepting state $q_a$ or rejects by entering the rejecting state $q_r$. We assume that $\delta$ is defined so that the input head never falls off an input tape out of the boundary symbols #, the storage tape head cannot write the blank symbol, and falls off the storage tape by moving left. $M$ halts when it enters state $q_a$ or $q_r$.

In this paper, we are concerned with 2-ptm's whose input tapes are restricted to square ones. Let $L : N \rightarrow N \cup \{0\}$ be a function, where $N$ denotes the set of all the positive integers. We say that a 2-ptm $M$ is $L(n)$ space-bounded if for each $n \geq 1$, and for each input tape $x$ with $l_1(x) = l_2(x) = n$, $M$ uses at most $L(n)$ cells of the storage tape. By 2-PTM$^*(L(n))$, we denote the class of square tapes recognized by $L(n)$ space-bounded 2-ptm's with error probability less than $\frac{1}{2}$ (whose input tapes are restricted to square ones). Especially, by 2-PTFA, we denote 2-PTM$^*(0)$, i.e., the class of sets of square tapes recognized by two-dimensional probabilistic finite automata [17] with error probability less than $\frac{1}{2}$.

A two-dimensional alternating finite automaton (2-afa) is a two-dimensional analogue of the alternating finite automaton [2] with the exception that the input tape moves left, right, up or down on the two-dimensional tape. See [9] for the formal definition of 2-afa's. By 2-AFA, we denote the class of sets accepted by 2-afa's whose input tapes are restricted to square ones. Throughout this paper, we assume that logarithms are base 2.

3. 2-AFA* versus 2-PTM$^*(L(n))$ with $L(n) = o(\log n)$

This section investigates a relationship between 2-AFA* and 2-PTM$^*(L(n))$ with $L(n) = o(\log n)$. We first give some preliminaries necessary for getting our desired result.

Let $M$ be a 2-ptm and the input alphabet of $M$. For each $m \geq 2$ and each $1 \leq n \leq m - 1$, an $(m, n)$-chunk over $\Sigma$ is a pattern as shown in Fig. 2, where $v_1 \in \Sigma^{(m-1) \times n}$ and $v_2 \in \Sigma^{m \times (m-n)}$. By $c_{m,n}(v_1, v_2)$, we denote the $(m, n)$-chunk as shown in Fig. 2. For any $(m, n)$-chunk $v$, we denote by $v(\#)$ the pattern obtained from $v$ by attaching the boundary symbols # to $v$ as shown in Fig. 3. Below, we assume without loss of generality that $M$ enters or exits the pattern $v(\#)$ only at the face designated by the bold line in Fig. 3. Thus, the number of the entrance points to $v(\#)$ (or the exit points from $v(\#)$) for $M$ is $n + 3$. We suppose that these entrance points (or exit points) are named $(2,0)^v, (2,1)^v, \ldots, (2,n)^v, (1,n+1)^v, (0, n+1)^v$ as shown in Fig. 4. Let $PT(v(\#))$ be the set of these entrance points (or exit points). To each cell of $v(\#)$, we assign a position as shown in Fig. 4. Let $PS(v(\#))$ be the set of all the positions of $v(\#)$. For each $n \geq 1$, an $n$-chunk over $\Sigma$ is a pattern in $\Sigma^{x \times n}$. For any $n$-chunk $u$, we denote by $u(\#)$ the pattern obtained from $u$ by attaching the boundary symbols # to $u$ as shown in Fig. 5. We again assume without loss of generality that $M$ enters or exits the pattern $u(\#)$ only at the face designated by the bold line in Fig. 5. The number of the entrance points to $u(\#)$ (or the exit points from $u(\#)$) for $M$ is again $n + 3$, and these entrance points (or exit points) are named $(2,0)^u, (2,1)^u, \ldots, (2,n)^u, (1,n+1)^u, (0, n+1)^u$ as shown in Fig. 5. Let $PT(u(\#))$ be the set of these entrance points (or exit points). For any $(m, n)$-chunk $v$ over $\Sigma$ and any $n$-chunk $u$ over $\Sigma$, let $v[u]$ be the tape in $\Sigma^{x \times n}$ consisting of $v$ and $u$ as shown in Fig. 6.

Let $M$ be a 2-ptm. A storage state of $M$ is a combination of the state of the finite control, the non-blank contents of the storage tape, and the storage tape head position. Let $q_a$ and $q_r$ be the accepting and rejecting states of $M$, respectively and $x$ be an $(m, n)$-chunk (or an $n$-chunk) over the input alphabet of $M$ ($m \geq 2, n \geq 1$). We define the chunk probabilities of $M$ on $x$ as follows. A starting condition for the chunk probability is a pair $(s, l)$, where $s$ is a storage state of $M$ and $l \in PT(x(\#))$; its intuitive meaning is "$M$ has just entered $x(\#)$ in storage state $s$ from entrance point $l$ of $x(\#)$". A stopping condition for the chunk probability is either:

(i) a pair $(s, l)$ as above, meaning that $M$ exits from $x(\#)$ in storage state $s$ at exit point $l$,
(ii) "Loop" meaning that the computation of $M$ loops forever within $x(\#)$,
(iii) "Accept" meaning that $M$ halts in the accepting state $q_a$ before exiting from $x(\#)$ at exit points of $x(\#)$, or
(iv) "Reject" meaning that $M$ halts in the rejecting state $q_r$ before exiting from $x(\#)$ at exit points of $x(\#)$.

For each starting condition $s$ and each stopping condition $\tau$, let $p(x, s, \tau)$ be the probability that stopping condition $\tau$ occurs given that $M$ is started in starting condition $s$ on an $(m, n)$-chunk (or $n$-chunk) $x$.

Computations of a 2-ptm are modeled by Markov chains [20] with finite state space, say $\{1, 2, \ldots, s\}$ for some $s$. A particular Markov chain is completely specified by its matrix $R = \{r_{ij}\}_{1 \leq i \leq j \leq s}$ of transition probabilities. If the Markov chain...
is in state $i$, then it next moves to state $j$ with probability $r_{ij}$. The chains we consider have the designated starting state, say, state 1, and some set $T_r$ of trapping states, so $r_{11} = 1$ for all $t \in T_r$. For $t \in T_r$, let $p^t[R, R]$ denote the probability that a Markov chain $R$ is trapped in state $t$ when started in state 1. The following lemma which bounds the effect of small changes in the transition probabilities of a Markov chain is used below.

Let $\beta > 1$. Say that two numbers $r$ and $r'$ are $\beta$-close if either (i) $r = r' = 0$ or (ii) $r > 0$, $r' > 0$ and $\beta^{-1} \leq \frac{r}{r'} \leq \beta$. Two Markov chains $R = (r_{ij})_{j=1}^{2}$ and $R' = (r'_{ij})_{j=1}^{2}$ are $\beta$-close if $r_{ij}$ and $r'_{ij}$ are $\beta$-close for all pairs $i, j$.

**Lemma 3.1** [3]. Let $R$ and $R'$ be two $s$-state Markov chains which are $\beta$-close, and let $t$ be a trapping state of both $R$ and $R'$. Then $p^t[R, R]$ and $p^t[R', R]$ are $\beta^3$-close.

**Theorem 3.1** There exists a set in 2-AFA, but not in 2-PTM$^*(L(n))$ for any $L(n) = o(\log n)$. 

**Proof.** Let $T_1 = \{e \in \{0, 1\}^{|t|} \mid |t| \geq 2\}$. Let $T_1$ act as follows. Given an input tape $x$ with $l_1(x) = l_2(x) \geq 2$, $M_1$ existentially chooses some row other than the top row, say the $i$-th row, of $x$. Then $M_1$ universally tries to check that, for each $j(1 \leq j \leq l_2(x))$, $x(i, j) = x(i, j)$. That is, on the $i$-th row and $j$-th column of $x(1 \leq j \leq l_2(x))$, $M_1$ enters a universal state to choose one of two further actions. One action is to pick up the symbol $x(i, j)$, move up with the symbol stored in the finite control, compare the stored symbol with the symbol $x(i, j)$, and enter an accepting state if both the symbols are identical. The other action is to continue to move right one tape cell (in order to pick up the symbol $x(i, j + 1)$ and compare it with the symbol $x(i, j + 1)$). It will be obvious that $M_1$ accepts $T$.

We next show that $T_1 \notin 2-PTM^*(L(n))$ with $L(n) = o(\log n)$. Suppose to the contrary that there exists a 2-ptm $M$ recognizing $T_1$ with error probability $\epsilon < \frac{1}{2}$. For large $n$, let

- $U(n) = \{e \mid |e| \geq 1, |t| \geq 2\}$
- $W(n) = \{0, 1\}^{2(n-1)}$, where $m_n = 2^{n} + 1$, and
- $V(n) = \{c_{(m_n, 0)}(w_1, w_2) \mid w_1 \in W(n) \land w_2 \in \{0\}^{m_n} \}$

We shall below consider the computations of $M$ on the input tapes of size-length $m_n$. For large $n$, let $C(n)$ be the set of all the storage states of $M$ using at most $L(m_n)$ storage tape cells, and let $c(n) = |C(n)|$. Then $c(n) = \beta c(m_n)$ for some constant $\beta$. Consider the chunk probabilities $p(v, \sigma, \tau)$ defined above. For each $(m_n, n)$-chunk $v$ in $V(n)$, there are a total of

$$d(n) = c(n) \times |PT(v(\#))| \times (c(n) \times |PT(v(\#))| + 3) = O(n^2 L(m_n))$$

chunk probabilities for some constant $t$. Fix some ordering of the pairs $(\sigma, \tau)$ starting and stopping conditions and let $p(v)$ be the vector of these probabilities according to this ordering.

We first show that if $v \in V(n)$ and if $p$ is a nonzero element of $p(v)$, then $p \geq 2^{-c(n)}a(n)$, where $a(n) = |PS(v(\#))| = O(m_n) = O(n^{2})$ for some constant $\beta$.

Form a Markov chain $K(v)$ with states of the form $(s, l)$, where $s$ is a storage state of $M$ and $l \in PS(v(\#)) \cup PT(v(\#))$. The chain state $(s, l)$ with $l \in PS(v(\#))$ corresponds to $M$ being in storage state $s$ scanning the symbol at position $l$ of $v(\#)$. Transition probabilities from such states are obtained from the transition probabilities of $M$ in the obvious way. For example, if the symbol at position $(i, j)$ of $v(\#)$ is $0$, and if $M$ in storage state $s$ reading a 0 can move its input head left and enter storage state $s'$, then the transition probability from state $(s', (i, j))$ to state $(s', (i, j - 1))$ is $1/2$. Chain states of the form $(s, (i, j))$ with $(i, j) \in PT(v(\#))$ are trap states of $K(v)$ which correspond to $M$ just having exited from $v(\#)$ in storage state $s$ at exact point $(i, j)$ of $v(\#)$. Now consider, for example, $p = p(v, \sigma, \tau)$, where $\sigma = (s, (i, j))$ and $\tau = (s', (k, l))$ with $(i, j) \in PT(v(\#))$. If $p > 0$, then there must be some path of nonzero probability in $K(v)$ from $(s, (i, j))$ to $(s', (k, l))$, and since $K(v)$ has at most $c(n)a(n)$ nontrapping states, there is such a path of length at most $c(n)a(n)$. Since $1/2$ is the smallest nonzero transition probability of $K(v)$, it follows that $p \geq 2^{-c(n)}a(n)$. If $\sigma = (s, (i, j))$ with $(i, j) \in PT(v(\#))$ and $\tau = Loop$, there must be a path of nonzero probability in $K(v)$ from state $(s, (i, j))$ to some state $(s', (i', j'))$ such that there is no path of nonzero probability from $(s', (i', j'))$ to any trap state of the form $(s', (k, l))$ with $(i', j') \in PT(v(\#))$. Again, if there is such a path, there is one of length at most $c(n)a(n)$. The remaining cases are similar.

For each $v = c_{(m_n, 0)}(w_1, w_2) \in V(n)$, let

$$contents(v) = \{u \in U(n) \mid u = w_1[(i, 1), (i, n)]\} \text{ for some } i(1 \leq i \leq 2n)$$

Divide $V(n)$ into contents-equivalence classes by making $v$ and $v'$ contents-equivalent if $contents(v) = contents(v')$. There are

$$contents(n) = \frac{2^n}{1} + \frac{2^n}{2} + \ldots + \frac{2^n}{2^n} = 2^{2n} - 1$$

contents-equivalence classes of $(m_n, n)$-chunks in $V(n)$. (Note that $contents(n)$ corresponds to the number of all the nonempty subsets of $U(n)$.) We denote by $CONTENTS(n)$ the set of all the representatives of these $contents(n)$ contents-equivalence classes. Of course, $|CONTENTS(n)| = contents(n)$. Divide $CONTENTS(n)$ into $M$-equivalence classes by making $v$ and $v'$ $M$-equivalent if $p(v)$ and $p(v')$ are zero in exactly the same coordinates. Let $E(n)$ be a largest $M$-equivalence class. Then we have

$$|E(n)| \geq contents(n)/2^{d(n)}$$

Let $d(n)$ be the number of nonzero coordinates of $p(v)$ for $v \in E(n)$. Let $\hat{p}(v)$ be the $d(n)$-dimensional vector of nonzero coordinates of $p(v)$. Note that $\hat{p}(v) \in [2^{-c(n)}a(n)]^{d(n)}$ for all $v \in E(n)$. Let $\log \hat{p}(v)$ be the componentwise $\log \hat{p}(v)$.
Then, \( \log \hat{p}(v) \in \left[ -c(n)a(n), 0 \right]^{e(n)} \). By dividing each coordinate interval \([-c(n)a(n), 0]^{e(n)} \) into subintervals of length \( \mu \), we divide the space \([-c(n)a(n), 0]^{e(n)} \) into at most \( (c(n)a(n)/\mu)^{e(n)} \) cells, each of size \( \mu \times \mu \times \ldots \times \mu \). We want to choose \( \mu \) so large enough that the number of cells is smaller than the size of \( E(n) \), that is,

\[
(c(n)a(n)/\mu)^{e(n)} < \frac{\text{contents}(n)}{2^{d(n)}} (\leq |E(n)|)
\]

(1)

Concretely, we choose \( \mu = 2^{-n} \). (From the assumption that \( L(n) = o(\log n) \), we have \( L(m_n) = o(\log m_n) \). Thus, \( L(m_n) = o(n) \). From this, by a simple calculation, we can easily see that for large \( n \), (1) holds for \( \mu = 2^{-n} \). Assuming (1), there must be two different \( (m_n, n) \)-chunks \( v, v' \in E(n) \) such that \( \log \hat{p}(v) \) and \( \log \hat{p}(v') \) belong to the same cell. Therefore, if \( p \) and \( p' \) are two nonzero probabilities in the same coordinate of \( p(v) \) and \( p(v') \), respectively, then

\[
|\log p - \log p'| \leq \mu.
\]

It follows that \( p \) and \( p' \) are \( 2^{n} \)-close. Therefore, \( p(v) \) and \( p(v') \) are componentwise \( 2^{n} \)-close.

For this \( v \) and \( v' \), we consider an \( n \)-chunk \( u \in \text{contents}(v) - \text{contents}(v') \). We describe two Markov chains, \( R \) and \( R' \), which model the computations of \( M \) on \( v[u] \) and \( v'[u] \), respectively. The state space of \( R \) is

\[
C(n) \times (PT(v[#]) \cup PT(u[#])) \cup \{ \text{Accept, Reject, Loop} \}.
\]

Thus the number of states of \( R \) is

\[
x = c(n)(n + 3n + 3) + 3 = 2c(n)(n + 3) + 3.
\]

The state \((s, (i, j)) \in c(n) \times PT(v[#])\) of \( R \) corresponds to \( M \) just having entered \( v[#] \) in storage state \( s \) from entrance point \((i, j)\) of \( v[#] \), and the state \((s', (k, l)) \in c(n) \times PT(u[#])\) of \( R' \) corresponds to \( M \) just having entered \( u[#] \) in storage state \( s' \) from entrance point \((k, l)\) of \( u[#] \). For convenience sake, we assume that \( M \) begins to read any input tape \( x \) in the initial storage state \( q_0 = (q_0, \lambda, 1) \), where \( q_0 \) is the initial state of \( M \), by entering \( x(1, 1) \) from the lower edge of the cell on which \( x(1, 1) \) is written. Thus, the starting state of \( R \) is Initial \((q_0, (2, 1)) \). The states Accept and Reject correspond to the computations halting in the accepting state and the rejecting state, respectively, and Loop means that \( M \) has entered an infinite loop. The transition probabilities of \( R \) are obtained from the chunk probabilities of \( M \) on \( v[#] \) and \( v'#[u]\). For example, the transition probability from \((s, (i, j))\) to \((s', (k, l))\) with \((i, j) \in PT(v[#])\) and \((k, l) \in PT(u[#])\) is just \( p(u, (s, (i, j)), (s', (k, l))) \), the transition probability from \((s', (k, l))\) to \((s, (i, j))\) with \((i, j) \in PT(v[#])\) and \((k, l) \in PT(u[#])\) is \( p(u, (s', (k, l)), (s, (i, j))) \), the transition probability from \((s, (i, j))\) to Accept is \( p(u, (s, (i, j)), \text{Accept}) \), and the transition probability from \((s', (k, l))\) to Accept is \( p(u, (s', (k, l)), \text{Accept}) \). The states Accept, Reject, and Loop are trap states. The chain \( R' \) is defined similarly, but using \( v'[u] \) in place of \( v[u] \).

Let \( \text{acc}(v[u]) \) (resp., \( \text{acc}(v'[u]) \)) be the probability that \( M \) accepts input \( v[u] \) (resp., \( \text{acc}(v'[u]) \)). Then, \( \text{acc}(v[u]) \) (resp., \( \text{acc}(v'[u]) \)) is exactly the probability that the Markov chain \( R \) (resp., \( R' \)) is trapped in state Accept when started in state Initial. From the fact that \( v[u] \) is in \( T_1 \), it follows that \( \text{acc}(v[u]) \geq 1 - \epsilon \). Since \( R \) and \( R' \) are \( 2^{n} \)-close, Lemma 3.1 implies that

\[
\frac{\text{acc}(v'[u])}{\text{acc}(v[u])} \geq 2^{-2\mu n}.
\]

\( 2^{-2\mu n} \) approaches 1 as \( n \) increases. Therefore, for large \( n \), we have

\[
\text{acc}(v'[u]) \geq 2^{-2\mu n} (1 - \epsilon) > \frac{1}{2},
\]

because \( \epsilon < \frac{1}{2} \). This is a contradiction, because \( v'[u] \notin T_1 \).

We conjecture that there is a set in 2-PFA*, but not in 2-AFA*. The candidate set is \( T_2 = \{ x \in \{0, 1\}^{n} | n \geq 2 \land \{ \text{the numbers of 0's and 1's in } x \text{ are the same} \} \) . By using the idea in [4], we can show that \( T_2 \) is in 2-PFA*. But, we have no proof of "\( T_2 \notin 2\text{-AFA}^* \)."

4. Space hierarchy between \( \log \log n \) and \( \log n \)

This section shows that there is an infinite space hierarchy for 2-ptm's with error probability less than \( \frac{1}{2} \) whose spaces are between \( \log \log n \) and \( \log n \).

A two-dimensional deterministic Turing machine (2-dtm) is a two-dimensional analogue of the two-way deterministic Turing machine [7], which has one read-only input tape and one semi-infinite read-write storage tape, with the exception that the input head moves left, right, up or down on the two-dimensional tape. The 2-dtm accepts an input tape \( x \) if it starts in the initial state with the input head on the upper-left hand corner of \( x \), and eventually enters an accepting state. See [9,16] for the formal definition of 2-dtms.

Let \( L(n) : N \to N \cup \{ 0 \} \) be a function. A 2-dtm \( M \) is strongly \( L(n) \) space-bounded if it uses at most \( L(n) \) cells of the storage tape for each \( n \geq 1 \) and each input tape \( x \) with \( l_1(x) = l_2(x) = n \). Let strong 2-DTM*(L(n)) be the class of sets of square tapes accepted by strongly \( L(n) \) space-bounded 2-dtm's.

A function \( L(n) : N \to N \cup \{ 0 \} \) is space-constructible by a two-dimensional Turing machine (2-tm) if there is a strongly \( L(n) \) space-bounded 2-dtm \( M \) such that for each \( n \geq 1 \), there exists some input tape \( x \) with \( l_1(x) = l_2(x) = n \) on which \( M \) halts after its storage tape head has marked off exactly \( L(n) \) cells of the storage tape. In this case, we say that \( M \) constructs the function \( L(n) \).
Let $\Sigma_1, \Sigma_2$ be finite sets of symbols. A projection is a mapping $\tau : \Sigma_1^{(2)} \to \Sigma_2^{(2)}$ which is obtained by extending the mapping $\tau : \Sigma_1 \to \Sigma_2$ as follows:

\[
\tau(x) = z' \iff 
\begin{align*}
(i) & \quad l_k(x) = l_k(x') \\
(ii) & \quad \tau(x(i,j)) = z'(i,j)
\end{align*}
\]

for each $k = 1, 2$, and for each $(i,j) (1 \leq i \leq l_1(x)$ and $1 \leq j \leq l_2(x))$.

### Theorem 4.1
If $L(n)$ is space-constructible by a 2-tm, log log $n < L(n) \leq \log n$, and $L'(n) = o(L(n))$, then, there exists a set in strong 2-DTM$^s(L(n))$, but not in 2-PTM$^s(L(n))$.

**Proof.** Let $L : N \to N$ be a function space-constructible by a two-dimensional Turing machine such that log log $n < L(n) \leq \log n$ ($n \geq 2$), and $M$ be a strongly $L(n)$ space-bounded 2-dtm which constructs the function $L$, and $T[L, M]$ be the following set, which depends on $L$ and $M$:

\[
T[L, M] = \{x \in (\Sigma \times \{0,1\})^{(2)}|n \geq 2 \quad l_1(x) = l_2(x) = n \& \exists \tau \leq L(n) \} ;
\]

when the tape $h_1(x)$ is presented to $M$, it uses $r$ cells of the storage tape and halts) & \exists (2 \leq \tau \leq n)[h_2(x([1,1], (1,1)) = h_2(x([1,1], (1,1)])],
\]

where $\Sigma$ is the input alphabet of $M$, and $h_1 (h_2)$ is the projection obtained by extending the mapping $h_1 : \Sigma \times \{0,1\} \to \Sigma (h_2 : \Sigma \times \{0,1\} \to \{0,1\})$ such that for any $c = (a, b) \in \Sigma \times \{0,1\}, h_1(c) = a \quad h_2(c) = b$.

We first show that $T[L, M] \in$ strong 2-DTM$^s(L(n))$. The set $T[L, M]$ is accepted by a strongly $L(n)$ space-bounded 2-dtm $M_1$ which acts as follows. When an input tape $x \in (\Sigma \times \{0,1\})^{(2)}$ with $l_1(x) = l_2(x) = n, n \geq 2$, is presented to $M_1$, $M_1$ directly simulates the action of $M$. If $M$ does not halt, then $M_1$ also does not halt, and will not accept $x$. If $M_1$ finds out that $M$ halts (in this case, note that $M_1$ has used at most $L(n)$ cells of the storage tape, because $M$ is a strongly $L(n)$ space-bounded), then $M_1$ checks by using the non-blank part of the storage tape that $h_2(x)$ is a desired form. $M_1$ enters an accepting state when such check is successful.

We next show that $T[L, M] \notin 2-PTM^s(L(n))$, where $L'(n) = o(L(n))$. For each $n \geq 2$, let $t(n) \in \Sigma^{(2)}$ be a fixed tape such that (i) $l_1(t(n)) = l_2(t(n)) = n$ and (ii) when $t(n)$ is presented to $M$, $M$ marks off exactly $L(n)$ cells of the storage tape and halts. (Note that for each $n \geq 2$, there exists such a tape $t(n)$, because $M$ constructs the function $L$.) Now, suppose that there exists a 2-ptm $M_2$ recognizing $T[L, M]$ with error probability $\epsilon < \frac{1}{2}$. We can derive a contradiction by using the same idea as in the proof of Theorem 3.1. The main difference is

(i) to replace

- "$U(n) =$ the set of all the $n$-chunks over $\{0,1\}^n$",
- "$W(n) = \{0,1\}^{(m_n-1)\times n}$, where $m_n = 2^n + 1$",
- "$V(n) = \{ch_{(m_{n},n)}(w_{1},w_{2})|w_{1} \in W(n) \& w_{2} \in \{0\}^{m_{n}\times (m_{n}-n)}\}$",
- "$c(n) = |C(n)| = b^{L(m_{n})}$, for some constant $b'$",
- "$d(n) = c(n) \times |PT(v(#))| \times (c(n) \times |PT(v(#))| + 3) = O(n^2 g(n_{2}))$",
- "$p \geq 2 - c(n) \ln(n)$, where $a(n) = |PS([v(#)])| = O(m_{n}^{2}) = O(e^n)$ for some constant $e^n$",
- "for each $v = ch_{(m_{n},n)}(w_{1},w_{2}) \in V(n)$, contents($v) = \{u \in U(n)[u = w_{1}(i,1), (i,n)]$ for some $i(1 \leq i \leq 2^n)\}$",
- "contents$[n] = \{1 \choose 1 + 2 \choose 1 + + 2^{2n} \times n - 1$ contents-equivalence classes of $(m_{n},n)$-chunks in $V(n)$",
- "$n = 2 - n$",
- "$n$-chunk $u \in contents(v) - contents(v')$", and
- "$z = c(n)(n + 3 + n + 3) + 3 = 2c(n)(n + 3 + 3)$",

in the proof of Theorem 3.1, with

- "$U(n) =$ the set of all the $L(n)$-chunks $u$ over $\Sigma \times \{0,1\}$ such that $h_1(u) = t(n)[(1,1), (1, L(n))])$",
- "$W(n) = \{u \in (\Sigma \times \{0,1\})^{(n-1)\times L(n)}|h_1(u) = t(n)[(2,1), (n, L(n))]\}$",
- "$V(n) = \{ch_{(n,L(n))}(w_{1},w_{2})|w_{1} \in W(n) \& w_{2}$ is a tape in $(\Sigma \times \{0\})^{n\times (n-L(n))}$ such that $h_1(w_{2}) = t(n)[(1, L(n+1)), (n,n)]\}$",
- "$c(n) = |C(n)| = b^{L(n)}$ for some constant $b'$",
- "$d(n) = c(n) \times |PT(v(#))| \times (c(n) \times |PT(v(#))| + 3) = O(L(n)^2 g(L(n))$ for some constant $t'$",
- "$p \geq 2 - c(n) \ln(n)$, where $a(n) = |PS([v(#)])| = O(m_{n}^{2})$",
- "for each $v = ch_{(n,L(n))}(w_{1},w_{2}) \in V(n)$, contents$[v) = \{u \in U(n)[u = w_{1}(i,1), (i,n)]$ for some $i(1 < i \leq n - 1)\}$",
- "$c(n) = \{2^{L(n)} + \ldots + 2^{L(n)} \}$ if $2^{L(n)} \geq n - 1$
- "$n = 2 - L(n)$",
- "$L(n)$-chunk $u \in contents(v) - contents(v')$", and
• "z = c(n)(L(n) + 3 + L(n) + 3) + 3 = 2c(n)(L(n) + 3) + 3",
respectively, and

(ii) to consider the computations on the input tapes of side-length n and on (n, L(n))-chunks, instead of considering the computations on the input tapes of side-length m_n and on (m_n, n)-chunks.

The details of the proof is left to the reader as an exercise. We note that by making a simple calculation, we can easily ascertain that

$$\left(\frac{c(n)a(n)}{\mu}\right)^{d(n)} \leq \frac{\text{contents}(n)}{2^d(n)} (\leq |E(n)|)$$

for large n and for our new c(n), a(n), d(n), µ, and contents(n), because log log n < L(n) ≤ log n and L'(n) = o(L(n)).

Since (log log n)^k, k ≥ 1, is space-constructible by a 2-tm (in fact, (log log n)^k is space-constructible by one-dimensional Turing machine [7]), it follows from Theorem 4.1 that the following corollary holds.

Corollary 4.1 For any integer k ≥ 1,

$$2\text{-PTM}((\log \log n)^k) \supsetneq 2\text{-PTM}((\log \log n)^{k+1}).$$

Remark 4.1 It is well-known [7] that, in the one-dimensional case, there exists no space-constructible function which grows more slowly than the order of log log n. On the other hand, Morita et al. [15] and Szepietowski [22] showed that the function log^{(k)}(n) (k ≥ 1), log^* n and log^{(k)} log^* n are all space-constructible by a two-dimensional Turing machine, where these functions are defined as follows:

$$\log^{(1)} n = \begin{cases} 0 & (n = 0) \\ \log_2 n & (n ≥ 1) \end{cases}$$

$$\log^{(k+1)} n = \log^{(1)}(\log^{(k)} n) \text{ for } k ≥ 1$$

$$\exp^* 0 = 1, \quad \exp^* (n+1) = 2^{\exp^* n}$$

$$\log^* n = \min\{x | \exp^x n ≥ 1\}$$

It is shown in [10,11,16] that for two-dimensional (deterministic, nondeterministic and alternating) Turing machines whose input tapes are restricted to square ones, log^{(k)} space-bounded machines are more powerful than log^{(k+1)} space-bounded machines (k ≥ 1). We conjecture that for each k ≥ 2, 2-PTM((log^{(k+1)} n) \supsetneq 2-PTM((log^{(k)} n)), but we have no proof of this conjecture.

5. Conclusion

We conclude this paper by giving the following open problems.

(1) For what L(n), is there a set in 2-PFA^k, but not accepted by any L(n) space-bounded two-dimensional alternating Turing machine?

(2) Is there an infinite space hierarchy for 2-ptm's with error probability ε < 1/2 whose spaces are below log log n?

It will be also interesting to investigate the relationship among the accepting powers of 2-ptm's with error probability ε < 1/2, 2-atm's with only universal states, and two-dimensional nondeterministic Turing machines [9]. We will discuss this topic in a forthcoming paper.

References


Figure 1: Two-dimensional probabilistic Turing machine.

Figure 2: \((m,n)\)-chunk.

Figure 3: \(v(\#)\).

Figure 4: An Illustration for \(v(\#)(v: (m,n)\text{-chunk})\).

Figure 5: An Illustration for \(u(\#)\).

Figure 6: \(v[u]\).