# A Note on the Ambiguity of Finite Automata

Thomas Y.-T. Lai* Shi-Chun Tsai†

## Abstract

*We give a new and simple proof of a key lemma on the rank of a matrix in Leung's paper [3], where he showed there exists a family of nondeterministic finite automata $\{A_n\}$, with $n$ states and exponential ambiguity, such that any smallest equivalent polynomially ambiguous finite automaton has $2^n - 1$ states.*

**Keyword:** nondeterministic automata, ambiguity, automata theory.

## 1 Introduction

Leung [3] resolved an open problem raised by Ravikumar and Ibarra [4] on the succinctness of representations relating to the types of ambiguity of finite automata. One of the main tools is finding the rank of some matrices corresponding to finite automata.

In this note we give a much simpler proof on the rank of a matrix, which is crucial for Leung's result. Matrix rank method has been used in many proofs of computational complexity problems [3, 1]. Not only for automata theory, it has also been widely used for communication complexity [5].

Given a non-deterministic finite automaton (NFA) $M$, the ambiguity of a string $w$ is the number of different accepting paths for $w$ in $M$. $M$ is unambiguous if the ambiguity of any string is either 0 or 1. Obviously, deterministic finite automata (DFA) is unambiguous. For any length $n$, the ambiguity of $M$ is the maximum ambiguity of all strings of length $n$.

It is clear that the ambiguity of any NFA is bounded by $k_0 k^n$, where $k_0$ is the number of initial states and $k$ is the number of states.

*Computer Science and Information Engineering Dept, National Chi-Nan University, 521 University Road, Pu-Li, Nan-Tou 545, Taiwan, Email: u6321024@ncnu.edu.tw.

†Computer Science and Information Engineering Dept, and Information Management Dept, National Chi-Nan University, 1 University Road, Pu-Li, Nan-Tou 545, Taiwan, Email: tsai@csie.ncnu.edu.tw. The work was supported in part by the National Science Council of Taiwan under contract NSC 89-2213-E-260-009.

There are five types of automata classified with the degree of ambiguity [3, 4]: deterministic finite automata(DFA), nondeterministic finite automata(NFA), unambiguous NFA(UFA), finitely ambiguous NFA(FNA) and polynomially ambiguous NFA(PNA). Following Leung's paper [3], let $C_1$ and $C_2$be any two of these five classes of automata.

- $C_1 \leq_p C_2$ ($C_1$can be polynomially converted to $C_2$): If there exists a polynomial $p$ such that for any finite automata in $C_1$ with $n$ states we can find an equivalent finite automata in $C_2$ with at most $p(n)$ states.

- $C_1 =_p C_2$ ($C_1$ is polynomially related to $C_2$): If $C_1 \leq_p C_2$ and $C_2 \leq_p C_1$.

- $C_1 \neq_p C_2$: If $C_1$ can be separated from $C_2$.

- $C_1 <_p C_2$ : If $C_1 \leq_p C_2$ and $C_1 \neq_p C_2$.

It is clear that DFA $\leq_p$ UFA, UFA $\leq_p$ FNA, FNA $\leq_p$ PNA, PNA $\leq_p$ NFA. Leung proved PNA $<$ NFA via a specially defined automata $A_n$ [3] and resolved an open problem proposed by Ravikumar and Ibarra [4]. Thus he established the relations: DFA $<_p$ UFA $<_p$ FNA $\leq_p$ PNA $<_p$ NFA. However, it is still open for FNA $<_p$ PNA. For any $n$, he considered automaton $A_n = (Q, \Sigma, \delta, \{q_1\}, \{q_1\})$, where $Q = \{q_1, \cdots, q_n\}$, $q_1$ is the only starting state and final state, $\Sigma = \{0, 1\}$ and $\delta(q_1, 0) = \{q_1, q_2\}$, $\delta(q_i, 0) = \{q_{i+1}\}$ for $2 \leq i \leq n-1$, $\delta(q_n, 0) = \{q_1\}$, $\delta(q_1, 1) = \emptyset$, and $\delta(q_i, 1) = \{q_i\}$ for $2 \leq i \leq n$.

It is not hard to see $L(A_n) = (0+(01^*)^{n-1}0)^*$. Along the way, he proved that the smallest DFA for $L(A_n)$ has $2^n$ states and the smallest UFA for $L(A_n)$ has $2^n - 1$ states. To prove the latter, the rank of a matrix related to $A_n$ is crucial. We give a simpler proof on proving the rank of the matrix.

We give the basic definitions in section 2. Section 3 presents the main result.

## 2 Preliminaries

In this note we assume that the reader is familiar with the basic terminologies in automata theory [2]. Consider the family of automata $A_n = (Q, \Sigma, \delta, \{q_1\}, \{q_1\})$ as defined above. For any $P \subseteq Q$, let $w_P \in \Sigma^*$ be $w_1 0 w_n 0 w_{n-1} 0 \cdots 0 w_1$ where $w_i = \epsilon$ if $q_i \in P$ and $w_i = 1$ otherwise; and let $u_P$ be $0^{n-1} w_P$.

**Lemma 1** *[3] For any $P \subseteq Q$, we have: (1) $\delta(P, w_P) = P$; (2) for any $q \in Q$, $\{q\} \subseteq \delta(q, w_P)$; (3) $\delta(Q - P, w_P) = \emptyset$.*

**Corollary 2** *For any $P \subseteq Q, \delta(q_1, u_P) = P$.*

**Corollary 3** *For any $P$ and $P' \subseteq Q$, $u_P w_{P'} \in prefix(L(A_n))$ iff $P \cap P' \neq \emptyset$.*

As in Leung's paper [3], let $M_n$ be a $2^n - 1 \times 2^n - 1$ matrix over the field of characteristic 2 with rows and columns indexed by the nonempty subsets of $Q$ such that $M_n(P, P') = 1$ if $u_p w'_p 0^{n-1} \in L_n$, and $M_n(P, P') = 0$ otherwise. By the above corollary, as observed by Leung, we know $M_n(P, P') = 1$ if $P \cap P' \neq \emptyset$ and 0 otherwise.

## 3 Main result

We give a new and simple proof of the following lemma, which is crucial in Leung's paper.

**Lemma 4** *The rank of $M_n$ is $2^n - 1$.*

**Proof.** By definition, we can index rows and columns of $M_n$ by $n$-bit binary numbers in increasing order such that any $n$-bit binary number $b_n b_{n-1} \cdots b_1$ corresponds to the nonempty subset $P \subseteq Q$ with the property that for any $1 \leq i \leq n$ , $q_i \in P$ iff $b_i = 1$. The indices range from binary number of value 1 to binary number of value $2^n - 1$. Let $a$ and $b$ be two $n$-bit binary numbers. Then $M_n(a, b) = 1$ if there is some $i$ such that $a_i = b_i = 1$, and 0 otherwise.The following matrix is for $n = 3$.

$$M_3 = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

Also note that if $a + b > 2^n - 1$, then by pigeon hole principle there exists some $i$ such that $a_i = b_i = 1$ and so $M_n(a, b) = 1$. For $a + b = 2^n - 1$, $M_n(a, b) = 0$, since $a$ and $b$ are complementary. Thus the entries on the auxiliary diagonal (from top right to bottom left) of $M_n$ are all 1. The entries to the right of the auxiliary diagonal are all also 1. The entries on the diagonal right above the auxiliary diagonal are all zero, since the sum of their row and column indices are exactly $2^n - 1$. By a sequence of proper row operations, we can transform $M_n$ into a triangular matrix with 0's on the lower right part and 1's on the auxiliary diagonal. It is clear that the rank of such triangular matrix is $2^n - 1$. More precisely, the row operations are: subtract row $i - 1$ from row $i$ of $M_n$ for $i = 2^n - 1$ down to 2. This completes the proof. **QED**

By the lemma, Leung proved the following results and eventually resolved an open queston proposed by Ravikumar and Ibarra, i.e., NFA is more powerful than PNA.

**Lemma 5** *[3] A smallest UFA recognizing $L(A_n)$ has $2^n - 1$ states.*

**Theorem 6** *[3] A smallest PNA recognizing $L(A_n)$ has $2^n - 1$ states.*

## 4 Conclusion

We prove the rank of $M_n$ in a much simpler way. This result is crucial in Leung's paper. However, the proof on the rank by Leung is complicated. Via a better observation on the structure of the matrix we shorten the proof significantly.

## References


[1] A. Condon, L. Hellerstein, S. Pottle and A. Wigderson, *On the power of finite automata with both nondeterministic and probabilistic states*, SIAM J. Comput., 27 (1998), pp. 739–762.

[2] J. Hopcroft and J. Ulmman, *Introduction to Automata Theory, Language and Computation*, Addison-Wesley, Reading, MA, 1979.

[3] H. Leung, *Separating exponentially ambiguous finite automata from polynomially ambiguous finite automata*, SIAM J. Comput., 27 (1998), pp. 1073–1082.

[4] B. Ravikumar and O. Ibarra, *Relating the type of ambiguity of finite automata to the succinctness of*

*their representation*, SIAM J. Comput., 18 (1989), pp.1263–1282.

[5] A. C. Yao, *Some complexity questons related to distributed computing*, in Proc. 11th Annual ACM Symposium on Theory of Computing, ACM, New York, 1979, pp. 209–213.