

GENERALIZED AUTOMATA AND STOCHASTIC LANGUAGES

PAAVO TURAKAINEN

1. Introduction. In a probabilistic automaton, as defined e.g. in [2] and [1], the initial vector as well as the transition matrices are stochastic, and the final vector consists of 0's and 1's only. This paper is concerned with the so-called generalized automata where the elements of the initial vector, of the final vector and of the matrices are allowed to be arbitrary real numbers. Our purpose is to prove that a language is accepted by a generalized automaton if and only if it is accepted by a probabilistic automaton, i.e., if and only if it is a stochastic language. As an application, we show that the mirror image of a stochastic language is a stochastic language.

2. Definitions and notations. The set of words, including the empty word Λ , over a finite alphabet I is denoted by $W(I)$. Subsets of $W(I)$ are called *languages*. The length of a word $P \in W(I)$ is denoted by $l(P)$.

For any matrix A , the notation A^T means the transpose of A .

DEFINITION. A *generalized automaton* over the alphabet I is an ordered quadruple $\mathfrak{G}\mathfrak{A} = (S, M, \pi_0, f_0)$ where $S = \{s_1, \dots, s_n\}$ is a finite nonempty set (the set of *states*), M is a mapping of I into the set of $n \times n$ matrices with real elements, π_0 is an n -dimensional row vector with real components (*initial vector*) and f_0 is an n -dimensional column vector with real components (*final vector*).

The domain of M is extended from I to $W(I)$ by defining

$$\begin{aligned} M(\Lambda) &= E_n \quad (n \times n \text{ identity matrix}), \\ M(x_1 x_2 \dots x_k) &= M(x_1) M(x_2) \dots M(x_k), \end{aligned}$$

where $k \geq 2$ and $x_i \in I$.

For any real number η , the language *accepted* by $\mathfrak{G}\mathfrak{A}$ with the *cut-point* η is defined by

$$L(\mathfrak{G}\mathfrak{A}, \eta) = \{P \in W(I) \mid \pi_0 M(P) f_0 > \eta\}.$$

A language L is called a $\mathfrak{G}\mathfrak{A}$ -*language* if and only if, for some $\mathfrak{G}\mathfrak{A}$ and η , $L = L(\mathfrak{G}\mathfrak{A}, \eta)$.

If π_0 is a stochastic vector, f_0 consists of 0's and 1's only and the matrices $M(x)$ ($x \in I$) are stochastic, then $\mathfrak{G}\mathfrak{A}$ is a *finite probabilistic automaton* and the language $L(\mathfrak{G}\mathfrak{A}, \eta)$ is a *stochastic language*.

Received by the editors May 23, 1968.

3. Lemmas. To show how, for any generalized automaton, a finite probabilistic automaton accepting the same language can be constructed, we begin with five lemmas.

LEMMA 1. *Every \mathcal{GA} -language L is accepted by a generalized automaton $\mathcal{GA}_1 = (S_1, M_1, \pi_1, f_1)$ where, for each $x \in I$, the row and column sums of $M_1(x)$ equal zero.*

PROOF. Let $L = L(\mathcal{GA}, \eta)$, where $\mathcal{GA} = (S, M, \pi_0, f_0)$ is an n -state generalized automaton. Clearly, for each $x \in I$, there exist real numbers $\alpha_1(x), \dots, \alpha_n(x), \beta_0(x), \dots, \beta_n(x)$ such that in the matrix

$$(1) \quad M_1(x) = \left[\begin{array}{c|ccc|c} 0 & 0 & \dots & 0 & 0 \\ \hline \alpha_1(x) & & & & 0 \\ \vdots & & M(x) & & \vdots \\ \alpha_n(x) & & & & 0 \\ \hline \beta_0(x) & \beta_1(x) & \dots & \beta_n(x) & 0 \end{array} \right]$$

every row and column sum equals 0. From this construction it follows that, for any nonempty word $P \in W(I)$ and for some real numbers $\alpha_1(P), \dots, \alpha_n(P), \beta_0(P), \dots, \beta_n(P)$, $M_1(P)$ is obtained from (1) by replacing x by P . Consequently, if we define $\pi_1 = (0, \pi_0, 0)$ and $f_1 = (0, f_0^T, 0)^T$ then, for the $(n+2)$ -state generalized automaton $\mathcal{GA}_1 = (S_1, M_1, \pi_1, f_1)$, the equation

$$\pi_1 M_1(P) f_1 = \pi_0 M(P) f_0$$

holds whenever $P \in W(I)$. This implies that $L = L(\mathcal{GA}_1, \eta)$, whence the lemma follows.

LEMMA 2. *Every \mathcal{GA} -language L is accepted by a generalized automaton $\mathcal{GA}_1 = (S_1, M_1, \pi_1, f_1)$ where, for each $x \in I$, the elements of $M_1(x)$ are nonnegative.*

PROOF. By Lemma 1, we may assume that $L = L(\mathcal{GA}, \eta)$ for an n -state generalized automaton $\mathcal{GA} = (S, M, \pi_0, f_0)$ where the row and column sums of the matrices $M(x)$ ($x \in I$) equal 0.

For any real number a , denote by $N(a)$ the $n \times n$ matrix whose elements equal a . Let $\delta > 0$ be so large that, for each $x \in I$, the elements of the matrix $M_2(x) = M(x) + N(\delta)$ are nonnegative. By the

assumption concerning the matrices $M(x)$ ($x \in I$), both $M(x)N(a)$ and $N(a)M(x)$ are zero matrices. This implies that, for any $x \in I$ and $y \in I$,

$$(2) \quad M_2(xy) = M_2(x)M_2(y) = M(xy) + N(n\delta^2).$$

It is easy to verify that the row and column sums of $M(xy)$ equal 0. Let $P \in W(I)$ be an arbitrary nonempty word. Proceeding inductively, we infer from (2) that

$$(3) \quad M_2(P) = M(P) + N(n^{l(P)-1}\delta^{l(P)}).$$

Let A be the 2×2 matrix whose rows equal $(0, 1)$. Consider the $(2n+2)$ -state generalized automaton $\mathfrak{G}\mathfrak{A}_1 = (S_1, M_1, \pi_1, f_1)$ where

$$\pi_1 = (\pi_0, \pi_0, \pi_0 f_0, 0), \quad f_1 = (f_0^T, -f_0^T, 1, 0)^T$$

and, for each $x \in I$,

$$M_1(x) = \begin{bmatrix} M_2(x) & 0 & 0 \\ 0 & N(\delta) & 0 \\ 0 & 0 & A \end{bmatrix}.$$

Consequently, for any nonempty word $P \in W(I)$,

$$(4) \quad M_1(P) = \begin{bmatrix} M_2(P) & 0 & 0 \\ 0 & N(n^{l(P)-1}\delta^{l(P)}) & 0 \\ 0 & 0 & A \end{bmatrix}.$$

Formulas (3) and (4) together with the definition of $M_1(\Lambda)$ now imply that, for any word $P \in W(I)$,

$$\pi_1 M_1(P) f_1 = \pi_0 M(P) f_0.$$

Thus $L = L(\mathfrak{G}\mathfrak{A}_1, \eta)$, whence the lemma follows.

LEMMA 3. Every $\mathfrak{G}\mathfrak{A}$ -language L is accepted by a generalized automaton $\mathfrak{G}\mathfrak{A}_1 = (S_1, M_1, \pi_1, f_1)$ where the matrices $M_1(x)$ ($x \in I$) are stochastic.

PROOF. By Lemma 2, we may assume that $L = L(\mathfrak{G}\mathfrak{A}, \eta)$ for an n -state generalized automaton $\mathfrak{G}\mathfrak{A} = (S, M, \pi_0, f_0)$ where the elements of the matrices $M(x)$ ($x \in I$) are nonnegative.

Let $\delta \geq 1$ be a real number larger than the largest row sum in the matrices $M(x)$ ($x \in I$). For each $x \in I$, there exist real numbers $\delta_i(x)$, $0 \leq \delta_i(x) \leq 1$, ($i = 1, \dots, n$) such that the matrix

$$M_2(x) = \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline \delta_1(x) & \\ \vdots & \delta^{-1}M(x) \\ \delta_n(x) & \end{array} \right]$$

is stochastic. From this construction it follows that, for any non-empty word $P \in W(I)$ and some real numbers $\delta_i(P)$ ($i=1, \dots, n$),

$$M_2(P) = \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline \delta_1(P) & \\ \vdots & \delta^{-l(P)}M(P) \\ \delta_n(P) & \end{array} \right].$$

Consider the $(n+3)$ -state generalized automaton

$$\mathfrak{U}_1 = (S_1, M_1, \pi_1, f_1)$$

where

$$\pi_1 = (0, \pi_0, \eta, 0), \quad f_1 = (0, f_0^T, -1, 0)^T$$

and, for each $x \in I$,

$$M_1(x) = \left[\begin{array}{c|cc} M_2(x) & & 0 \\ \hline & \delta^{-1} & 1 - \delta^{-1} \\ 0 & 0 & 1 \end{array} \right].$$

Consequently, for any nonempty word $P \in W(I)$,

$$M_1(P) = \left[\begin{array}{c|cc} M_2(P) & & 0 \\ \hline & \delta^{-l(P)} & 1 - \delta^{-l(P)} \\ 0 & 0 & 1 \end{array} \right].$$

We now conclude that, for any word $P \in W(I)$,

$$\pi_1 M_1(P) f_1 = \delta^{-l(P)} (\pi_0 M(P) f_0 - \eta).$$

Thus $L = L(\mathfrak{G}\mathfrak{A}_1, 0)$, whence the lemma follows.

LEMMA 4. *Every $\mathfrak{G}\mathfrak{A}$ -language L is accepted by a generalized automaton $\mathfrak{G}\mathfrak{A}_1 = (S_1, M_1, \pi_1, f_1)$ where the initial vector π_1 and the matrices $M_1(x)$ ($x \in I$) are stochastic.*

PROOF. By Lemma 3, we may assume that $L = L(\mathfrak{G}\mathfrak{A}, \eta)$ for an n -state generalized automaton $\mathfrak{G}\mathfrak{A} = (S, M, \pi_0, f_0)$ where the matrices $M(x)$ ($x \in I$) are stochastic.

Let $\delta > 0$ be so large that, for $\rho_0 = (p_1, \dots, p_n)$, the components of the vector $(p_1 + \delta, \dots, p_n + \delta)$ are positive. Denote $p = p_1 + \dots + p_n + 2n\delta$. Thus $p > 0$. Consider the $2n$ -state generalized automaton $\mathfrak{G}\mathfrak{A}_1 = (S_1, M_1, \pi_1, f_1)$ where

$$\pi_1 = p^{-1}(p_1 + \delta, \dots, p_n + \delta, \delta, \dots, \delta), \quad f_1 = \begin{bmatrix} f_0 \\ -f_0 \end{bmatrix}$$

and, for each $x \in I$,

$$M_1(x) = \begin{bmatrix} M(x) & 0 \\ 0 & M(x) \end{bmatrix}.$$

We conclude that, for any word $P \in W(I)$,

$$\pi_1 M_1(P) f_1 = p^{-1} \pi_0 M(P) f_0.$$

Thus $L = L(\mathfrak{G}\mathfrak{A}_1, \eta/p)$.

LEMMA 5. *Assume that $L = L(\mathfrak{G}\mathfrak{A}, \eta)$ where $\mathfrak{G}\mathfrak{A} = (S, M, \pi_0, f_0)$ is a generalized automaton such that the initial vector π_0 and the matrices $M(x)$ ($x \in I$) are stochastic. For any real number c , $L = L(\mathfrak{G}\mathfrak{A}_1, \eta + c)$ where $\mathfrak{G}\mathfrak{A}_1 = (S, M, \pi_0, f_1)$ and $f_1 = f_0 + (c, \dots, c)^T$.*

PROOF. For any word $P \in W(I)$,

$$\pi_0 M(P) f_1 = \pi_0 M(P) f_0 + c,$$

because $\pi_0 M(P)$ is a stochastic vector. This proves the lemma.

4. Equivalence theorem and an application. Using Lemmas 4 and 5, we now establish the following

THEOREM 1. *A language L is accepted by a generalized automaton if and only if it is accepted by a finite probabilistic automaton, i.e., if and only if it is a stochastic language.*

PROOF. The "if"-part is immediate. To prove the "only if"-part, we may assume, by Lemma 4, that $L = L(\mathcal{G}\mathfrak{A}, \eta)$ for an n -state generalized automaton $\mathcal{G}\mathfrak{A} = (\{s_1, \dots, s_n\}, M, \pi_0, f_0)$ where the initial vector $\pi_0 = (p_1, \dots, p_n)$ and the matrices $M(x)$ ($x \in I$) are stochastic. By choosing c large enough in Lemma 5, we may assume that the components of the final vector $f_0 = (q_1, \dots, q_n)^T$ are positive. Denote $q = q_1 + \dots + q_n$. Thus q_1, \dots, q_n and q are positive. Let $q'_i = q_i/q$ for each $i = 1, \dots, n$. Consider the n^2 -state probabilistic automaton

$$\mathcal{G}\mathfrak{A}_1 = (\{s_1, \dots, s_n, \dots, s_{n^2}\}, M_1, \pi_1, f_1)$$

defined as follows. Let

$$\pi_1 = n^{-1}(\pi_0, \dots, \pi_0) \quad (\pi_0 \text{ occurs } n \text{ times})$$

and f_1 an n^2 -dimensional column vector whose i th component equals 1 if $s_i \in \{s_1, s_{n+2}, s_{2n+3}, \dots, s_{n^2}\}$ and equals 0 otherwise. For each $x \in I$, we define

$$(5) \quad M_1(x) = \begin{bmatrix} q'_1 M(x) & q'_2 M(x) & \dots & q'_n M(x) \\ q'_1 M(x) & q'_2 M(x) & \dots & q'_n M(x) \\ \vdots & \vdots & \ddots & \vdots \\ q'_1 M(x) & q'_2 M(x) & \dots & q'_n M(x) \end{bmatrix}.$$

Let $P \in W(I)$ be an arbitrary nonempty word. From the construction of the matrices $M_1(x)$ it follows that (5) holds if x is replaced by the word P . This implies that

$$(6) \quad \pi_1 M_1(P) f_1 = \sum_{i=1}^n \sum_{j=1}^n p_i q'_j p_{ij}(P)$$

where $p_{ij}(P)$ denotes the (i, j) th element of the matrix $M(P)$. On the other hand,

$$(7) \quad \pi_0 M(P) f_0 = \sum_{i=1}^n \sum_{j=1}^n p_i q_j p_{ij}(P).$$

Combining (6) and (7), we obtain

$$\pi_1 M_1(P) f_1 = q^{-1} \pi_0 M(P) f_0.$$

This implies that the language $L(\mathcal{G}\mathfrak{A}_1, \eta/q)$ contains exactly the same nonempty words as the language $L(\mathcal{G}\mathfrak{A}, \eta)$. Since the union and the intersection of a stochastic language and a regular language are both stochastic languages, it follows that $L(\mathcal{G}\mathfrak{A}, \eta)$ is a stochastic language. This proves the theorem.

As an application of Theorem 1, we now establish the following

theorem, where the mirror image of a language L , in symbols $\text{mi}(L)$, means the language obtained from L by writing all words backwards.

THEOREM 2. *A language L is stochastic if and only if the mirror image of L is stochastic.*

PROOF. Since $\text{mi}(\text{mi}(L)) = L$, it suffices to prove that if L is a stochastic language, so is $\text{mi}(L)$.

Let $L = L(\mathcal{U}, \eta)$, where $\mathcal{U} = (S, M, \pi_0, f_0)$ is a finite probabilistic automaton. Consider the generalized automaton $\mathcal{U}_1 = (S, M_1, \pi_1, f_1)$ where $\pi_1 = f_0^T$, $f_1 = \pi_0^T$ and, for each $x \in I$, $M_1(x) = M(x)^T$. For each word $P \in W(I)$, we now obtain

$$\pi_1 M_1(\text{mi}(P)) f_1 = \pi_0 M(P) f_0,$$

which implies that $L(\mathcal{U}_1, \eta) = \text{mi}(L)$. Theorem 2 follows now from Theorem 1.

REFERENCES

1. A. Paz, *Some aspects of probabilistic automata*, Information and Control 9 (1966), 26–60.
2. M. O. Rabin, *Probabilistic automata*, Information and Control 6 (1963), 230–245.

UNIVERSITY OF TURKU, FINLAND