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 $\Lambda$, 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 $\mathcal{A} = (S, M, \pi_0, f_0)$ where $S = \{s_1, \ldots, s_n\}$ 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, $\pi_0$ 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

$$M(\Lambda) = E_n \quad (n \times n \text{ identity matrix}),$$

$$M(x_1x_2 \cdots x_k) = M(x_1)M(x_2) \cdots M(x_k),$$

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

For any real number $\eta$, the language accepted by $\mathcal{A}$ with the cut-point $\eta$ is defined by

$$L(\mathcal{A}, \eta) = \{P \in W(I) \mid \pi_0M(P)f_0 > \eta\}.$$

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

If $\pi_0$ 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 $\mathcal{A}$ is a finite probabilistic automaton and the language $L(\mathcal{A}, \eta)$ is a stochastic language.

Received by the editors May 23, 1968.

303
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 Ø-language $L$ is accepted by a generalized automaton $\mathfrak{A}_1 = (S_1, M_1, \pi_1, f_1)$ where, for each $x \in I$, the row and column sums of $M_1(x)$ equal zero.

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

$$
M_1(x) = \begin{bmatrix}
0 & 0 & \cdots & 0 & 0 \\
\alpha_1(x) & \cdots & \cdots & \cdots & 0 \\
\vdots & & & & \vdots \\
\alpha_n(x) & \cdots & \cdots & \cdots & 0 \\
\beta_0(x) & \beta_1(x) & \cdots & \beta_n(x) & 0
\end{bmatrix}
$$

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 $\alpha_1(P), \cdots, \alpha_n(P), \beta_0(P), \cdots, \beta_n(P), M_1(P)$ is obtained from (1) by replacing $x$ by $P$. Consequently, if we define $\pi_1 = (0, \pi_0, 0)$ and $f_1 = (0, f_0^T, 0)^\top$ then, for the $(n+2)$-state generalized automaton $\mathfrak{A}_1 = (S_1, M_1, \pi_1, f_1)$, the equation

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

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

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

**Proof.** By Lemma 1, we may assume that $L = L(\mathfrak{A}, \eta)$ for an $n$-state generalized automaton $\mathfrak{A} = (S, M, \pi_0, f_0)$ 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 $M_1(x) = M(x) + N(\delta)$ 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 $y \in I$,

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

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

$$M_2(P) = M(P) + N(n^{i(P)} - i^{i(P)}).$$

Let $A$ be the $2 \times 2$ matrix whose rows equal $(0, 1)$. Consider the $(2n+2)$-state generalized automaton $\mathcal{G}_2 = (S_1, M_{x_1}, \pi_1, f_1)$ where

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

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 $P \in W(I)$,

$$M_1(P) = \begin{bmatrix} M_2(P) & 0 & 0 \\ 0 & N(n^{i(P)} - i^{i(P)}) & 0 \\ 0 & 0 & A \end{bmatrix}.$$ 

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

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

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

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

**Proof.** By Lemma 2, we may assume that $L = L(\mathcal{G}_1, \eta)$ for an $n$-state generalized automaton $\mathcal{G}_1 = (S, M, \pi_0, f_0)$ 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 $\delta_i(x)$, $0 \leq \delta_i(x) \leq 1$, ($i = 1, \ldots, n$) such that the matrix
is stochastic. From this construction it follows that, for any non-empty word $P \in W(I)$ and some real numbers $\delta_i(P)$ ($i=1, \ldots, n$),

$$M_2(P) = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
\delta_1(P) & \vdots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
\delta_n(P) & \end{bmatrix}.$$

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

$$\mathcal{A}_1 = (S_1, M_1, \pi_1, f_1)$$

where

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

and, for each $x \in I$,

$$M_1(x) = \begin{bmatrix}
M_2(x) & 0 \\
\delta^{-1} & 1 - \delta^{-1} \\
0 & 0 & 1 \\
\end{bmatrix}.$$

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

$$M_1(P) = \begin{bmatrix}
M_2(P) & 0 \\
\delta^{-1}(P) & 1 - \delta^{-1}(P) \\
0 & 0 & 1 \\
\end{bmatrix}.$$
We now conclude that, for any word $P \in W(I)$,

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

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

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

**Proof.** By Lemma 3, we may assume that $L = L(\mathcal{A}_1, \eta)$ for an $n$-state generalized automaton $\mathcal{A}_1 = (S, M, \pi_0, f_0)$ where the matrices $M(x) \ (x \in I)$ are stochastic.

Let $\delta > 0$ be so large that, for $\pi_0 = (\pi_1, \ldots, \pi_n)$, the components of the vector $(\pi_1 + \delta, \ldots, \pi_n + \delta)$ are positive. Denote $p = \pi_1 + \cdots + \pi_n + 2n \delta$. Thus $p > 0$. Consider the $2n$-state generalized automaton $\mathcal{A}_2 = (S_2, M_2, \pi_2, f_2)$ where

$$\pi_2 = p^{-1}(\pi_1 + \delta, \ldots, \pi_n + \delta), \quad f_2 = \begin{bmatrix} f_0 & \eta \end{bmatrix}$$

and, for each $x \in I$,

$$M_2(x) = \begin{bmatrix} M(x) & 0 \\ 0 & M(x) \end{bmatrix}.$$ 

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 $L = L(\mathcal{A}_2, \eta/p)$.

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

**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 \( L = L(\mathfrak{A}, \eta) \) for an \( n \)-state generalized automaton \( \mathfrak{A} = (\{ s_1, \ldots, s_n \}, M, \pi_0, f_0) \) where the initial vector \( \pi_0 = (p_1, \ldots, p_n) \) 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 \( f_0 = (q_1, \ldots, q_n)^T \) are positive. Denote \( q = q_1 + \cdots + q_n \). Thus \( q_1, \ldots, q_n \) and \( q \) are positive. Let \( q_i = q_i / q \) for each \( i = 1, \ldots, n \). Consider the \( n^2 \)-state probabilistic automaton

\[
\mathfrak{A}_1 = (\{ s_1, \ldots, s_n, \ldots, s_n^2 \}, M_1, \pi_1, f_1)
\]

defined as follows. Let

\[
\pi_1 = n^{-1}(\pi_0, \ldots, \pi_0) \quad (\pi_0 \text{ occurs } n \text{ times})
\]

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

\[
M_1(x) = \begin{bmatrix}
q_i M(x) & q_i M(x) & \cdots & q_i M(x) \\
q_i M(x) & q_i M(x) & \cdots & q_i M(x) \\
q_i M(x) & q_i M(x) & \cdots & q_i M(x)
\end{bmatrix}.
\]

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

\[
\pi_1 M_1(P) f_1 = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} q_i P_{ij}(P)
\]

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

\[
\pi_0 M(P) f_0 = \sum_{i=1}^{n} \sum_{j=1}^{n} p_{ij} p_{ij}(P).
\]

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

As an application of Theorem 1, we now establish the following
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 \( L = L(\mathfrak{A}, \eta) \), where \( \mathfrak{A} = (S, M, \tau_0, f_0) \) is a finite probabilistic automaton. Consider the generalized automaton \( \mathfrak{A}_1 = (S, M_1, \tau_1, f_1) \) where \( \tau_1 = f_0^T \), \( f_1 = \tau_0^T \) and, for each \( x \in I \), \( M_1(x) = M(x)^T \). For each word \( P \in W(I) \), we now obtain

\[
\tau_1 M_1(\text{mi}(P))f_1 = \tau_0 M(P)f_0,
\]

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

References


University of Turku, Finland