Cascade Products of Stochastic Automata

MERVE NUR CAKIR, MEHWISH SALEEMI, KARL-HEINZ ZIMMERMANN
Department of Computer Engineering
Hamburg University of Technology
20171 Hamburg
GERMANY

Abstract: - Stochastic Moore automata have in opposition to stochastic Mealy automata the same capabilities as general stochastic automata, but have the advantage that they are easier to access than their pure stochastic counterparts. Cascade decomposition of automata leads to a loop-free partitioning and in this way contributes to the analysis of automata. This paper shows that stochastic Moore automata can be decomposed into cascade products of stochastic Moore automata under mild conditions.

Key-Words: - Semiautomaton, stochastic automaton, Mealy automaton, Moore automaton, cascade product, mutual independence

1 Introduction

The theory of discrete stochastic systems has been initiated by the work of Shannon [1] and von Neumann [2]. While Shannon has considered memoryless communication channels and their generalization by introducing states, von Neumann has studied the synthesis of reliable systems from unreliable components. The fundamental work of Rabin and Scott [3] about deterministic finite-state automata has led to two generalizations. First, the generalization of transition functions to conditional distributions studied by Carlyle [4] and Starke [5]. This in turn yields a generalization of discrete-time Markov chains in which the chains are governed by more than one transition probability matrix [6]. Second, the generalization of regular sets by introducing stochastic automata as described by Rabin [7].

By the work of Turakainen [8], stochastic acceptors can be equivalently viewed as generalized automata in which the "probability" is neglected. This leads to a more accessible approach of stochastic acceptors [9, 10]. On the other hand, the class of nondeterministic automata [11] can be generalized to monoidal automata, where the input alphabet corresponds to an arbitrary monoid instead of a free monoid [12, 13, 14]. This leads to socalled monoidal automata whose corresponding class of monoidal languages is closed under a smaller set of operations when compared with the class of regular languages.

The famous result of Krohn-Rhodes states that deterministic finite-state automata can be decomposed into elementary components which correspond to permutation and reset automata connected by cascade products [13, 15, 16]. In the following, the cascade product defined for deterministic finite-state automata will be adapted to stochastic automata.

Mealy and Moore automata form a class of deterministic finite-state input-output automata which are of theoretical interest but also utilized in hardware design [17, 18, 19]. There are stochastic variants of Mealy and Moore automata, where its known that the class of stochastic Mealy automata forms a subclass of the stochastic Moore automata and the class of stochastic Moore automata already corresponds to the full class of stochastic automata [9, 10, 20]. Due to their structure, stochastic Moore automata are easier accessible than general stochastic automata.

In this paper, stochastic automata are considered in section 2 and stochastic variants of Mealy and Moore automata in section 3. In the main section 4, a general notion of cascade product of stochastic automata following the lines of deterministic finite-state automata in the Krohn-Rhodes theory is introduced. It will be shown that the cascade product of a stochastic Moore automaton yields a stochastic Moore automaton and each stochastic Moore automaton has a cascade decomposition into stochastic Moore automata under mild conditions.

2 Stochastic Automata

Stochastic automata are abstract finite-state machines with input and output behavior [9, 10, 12, 21]. A stochastic automaton can be viewed as an extension of a nondeterministic finite-state automaton with probabilistic transitions.

Formally, a *stochastic automaton* (SA) is a quadruple $A = (S, \Sigma, \Omega, p)$, where S is a nonempty

finite set of *states*, Σ is an alphabet of *input symbols*, Ω is an alphabet of *output symbols*, and for each pair $(a, s) \in \Sigma \times S$, $p(\cdot, \cdot \mid a, s)$ is a conditional probability distribution on $\Omega \times S$.

Given a conditional probability distribution $p(\cdot, \cdot \mid a, s)$ on $\Omega \times S$, we define a conditional probability distribution $\hat{p}(\cdot, \cdot \mid x, s)$ with $x \in \Sigma^*$ and $s \in S$ on $\Omega^* \times S$ recursively as follows.

- For all $s, s' \in S$,

$$\hat{p}(\epsilon, s' \mid \epsilon, s) = \begin{cases} 1 & \text{if } s = s', \\ 0 & \text{if } s \neq s', \end{cases} \quad (1)$$

where ϵ denotes both the empty word in Σ^* and Ω^*.

- For all $s, s' \in S$, $x \in \Sigma^*$, and $y \in \Omega^*$ with $|x| \neq |y|$,

$$\hat{p}(y, s' \mid x, s) = 0. \quad (2)$$

- For all $s, s' \in S$, $a \in \Sigma$, $x \in \Sigma^*$, $b \in \Omega$, and $y \in \Omega^*$,

$$\hat{p}(yb, s' \mid xa, s) = \sum_{t \in S} \hat{p}(y, t \mid x, s) \cdot p(b, s' \mid a, t). \quad (3)$$

Then for each pair $(x, s) \in \Sigma^* \times S$, $\hat{p}(\cdot, \cdot \mid x, s)$ is a conditional probability distribution on $\Omega^* \times S$.

Note that the measures p and $\hat{p}$ coincide on the set $\Omega \times S \times \Sigma \times S$ if we put $x = y = \epsilon$ in (3). Therefore, we will write p instead of $\hat{p}$.

As a generalization of (3), for all $x, x' \in \Sigma^*$, $y, y' \in \Omega^*$, and $s, s' \in S$ with $|x| = |y|$,

$$p(yy', s' \mid xx', s) = \sum_{t \in S} p(y, t \mid x, s) \cdot p(y', s' \mid x', t). \quad (4)$$

A stochastic automaton works serially and synchronously. It reads an input word symbol by symbol and after reading an input symbol it emits an output symbol and transits into another state. More precisely, if the automaton starts in state s and reads the word x, then with probability $\hat{p}(y, s' \mid x, s)$ it will end in state s' emitting the word y by taking all intermediate states into account.

The behavior of a stochastic automaton can be described by probability matrices. To this end, let A be a stochastic automaton with state set $S = \{s_1, \ldots, s_n\}$. For each pair of input and output symbols $a \in \Sigma$ and $b \in \Omega$, put $p_{ij}(b \mid a) = p(b, s_j \mid a, s_i)$ for all $1 \leq i, j \leq n$ and define the real-valued $n \times n$ matrix

$$P(b \mid a) = (p_{ij}(b \mid a))_{1 \leq i,j \leq n}. \quad (5)$$

Note that the matrix $P(b \mid a)$ is *substochastic*, i.e., it is a square matrix with nonnegative entries and each row adds up to at most 1. The elements of $P(b \mid a)$ provide the transition probabilities between the states if the symbol a is read and the symbol b is emitted. This definition can be extended to strings of input and output symbols. For this, note that by (1) we have

$$P(\epsilon \mid \epsilon) = I_n, \quad (6)$$

where I_n is the $n \times n$ unit matrix. Moreover, if $x \in \Sigma^*$ and $y \in \Omega^*$ with $|x| \neq |y|$, then by (2) we have

$$P(x \mid y) = O_n, \quad (7)$$

where O_n is the $n \times n$ zero matrix. Furthermore, if $a \in \Sigma$, $x \in \Sigma^*$, $b \in \Omega$, and $y \in \Omega^*$, then by (3) we have

$$P(yb \mid xa) = P(y \mid x) \cdot P(b \mid a). \quad (8)$$

More generally, by (4) and the associativity of matrix multiplication, for all $x, x' \in \Sigma^*$ and $y, y' \in \Omega^*$ with $|x| = |y|$,

$$P(yy' \mid xx') = P(y \mid x) \cdot P(y' \mid x'). \quad (9)$$

Summing up, each stochastic automaton A is uniquely characterized by the collection of substochastic matrices $\{P(b \mid a) \mid a \in \Sigma, b \in \Omega\}$. For each input symbol $a \in \Sigma$, the matrix $P(a) = \sum_{b \in \Omega} P(b \mid a)$ is stochastic. More generally, for each input word $x \in \Sigma^*$, the matrix $P(x) = \sum_{y \in \Omega^*} P(y \mid x)$ is stochastic.

Example 1. *Consider the stochastic automaton $A = (\{s_1, s_2\}, \{a\}, \{b\}, p)$ with conditional probabilities*

$$p(b, s_1 \mid a, s_1) = \frac{2}{3}, \; p(b, s_2 \mid a, s_1) = \frac{1}{3},$$
$$\text{and } p(b, s_2 \mid a, s_2) = 1.$$

The automaton is given by the state diagram in Fig. 1. The corresponding substochastic matrix is

$$P(a) = P(b \mid a) = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ 0 & 1 \end{pmatrix}.$$

Thus for each integer $k \geq 1$,

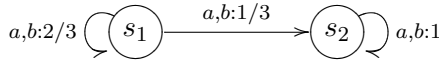$$P(a^k) = \begin{pmatrix} \frac{2^k}{3^k} & \frac{3^k - 2^k}{3^k} \\ 0 & 1 \end{pmatrix}.$$

$\diamond$

Figure 1: State diagram of A.

3 Stochastic Mealy and Moore Automata

Mealy and Moore automata are prominent classes of automata [17, 18, 19]. In a Mealy automaton the output is determined both by the current state and the current input, while in a Moore automaton the output is defined by the next state. In the following, stochastic versions of these automata types are presented [9, 10, 20].

A stochastic automaton $A = (S, \Sigma, \Omega, p)$ is called *stochastic Mealy automaton* if there are conditional probabilities $p_1(\cdot \mid a, s)$ and $p_2(\cdot \mid a, s)$ over Ω and S, respectively, such that for all $a \in \Sigma$, $b \in \Omega$ and $s, s' \in S$,

$$p(b, s' \mid a, s) = p_1(b \mid a, s) \cdot p_2(s' \mid a, s). \quad (10)$$

Stochastic independence of state transition and output emission makes stochastic Mealy automata a restricted class of stochastic automata.

A stochastic automaton A is *state-determined* if there is a mapping $\delta : \Sigma \times S \to S$ such that for all $a \in \Sigma$ and $s \in S$,

$$\sum_{b \in \Omega} p(b, \delta(a, s) \mid a, s) = 1. \quad (11)$$

Each state-determined stochastic automaton A is a stochastic Mealy automaton.

A stochastic automaton A is *output-determined* if there is a mapping $\lambda : \Sigma \times S \to \Omega$ such that for all $a \in \Sigma$ and $s \in S$,

$$\sum_{s' \in S} p(\lambda(a, s), s' \mid a, s) = 1. \quad (12)$$

Each output-determined stochastic automaton A is a stochastic Mealy automaton.

Example 2. *Take the stochastic Mealy automaton*

$$A = (\{s_1, s_2\}, \{a\}, \{b, c\}, p)$$

with probabilities

	$p_1(b \mid a, \cdot)$	$p_1(c \mid a, \cdot)$
s_1	$\frac{1}{2}$	$\frac{1}{2}$
s_2	$\frac{1}{4}$	$\frac{3}{4}$

	$p_2(s_1 \mid a, \cdot)$	$p_2(s_2 \mid a, \cdot)$
s_1	$\frac{1}{3}$	$\frac{2}{3}$
s_2	$\frac{1}{5}$	$\frac{4}{5}$

Then

$$P(b \mid a) = \begin{pmatrix} \frac{1}{6} & \frac{1}{3} \\ \frac{1}{20} & \frac{1}{5} \end{pmatrix} \quad and \quad P(c \mid a) = \begin{pmatrix} \frac{1}{6} & \frac{1}{3} \\ \frac{3}{20} & \frac{3}{5} \end{pmatrix}$$

and therefore

$$P(a) = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{5} & \frac{4}{5} \end{pmatrix}.$$

$\diamondsuit$

A stochastic automaton $A = (S, \Sigma, \Omega, p)$ is called *stochastic Moore automaton* if there are conditional probabilities $\mu(\cdot \mid s)$ and $p'(\cdot \mid a, s)$ over Ω and S, respectively, such that for all $a \in \Sigma$, $b \in \Omega$ and $s, s' \in S$,

$$p(b, s' \mid a, s) = \mu(b \mid s') \cdot p'(s' \mid a, s). \quad (13)$$

In opposition to stochastic Mealy automata, the output emission in Moore automata depends only on the state transition. The following result shows that stochastic Moore automata already provide the most general type of stochastic automata [9, 10].

Theorem 1. *For each stochastic automaton $A = (S_A, \Sigma, \Omega, p_A)$ there is a stochastic Moore automaton $B = (S_B, \Sigma, \Omega, p_B)$ such that A and B are S-equivalent.*

Proof. First, put $S_B = \Omega \times S_A$ and

$$p_B(b_2, (b_1, s_1) \mid a, (b_0, s_0)) = \begin{cases} p_A(b_1, s_1 \mid a, s_0) & \text{if } b_1 = b_2, \\ 0 & \text{otherwise} \end{cases}$$

for all $a \in \Sigma$, $b_2 \in \Omega$, and $(b_1, s_1), (b_0, s_0) \in S_B$.

Moreover, for all $a \in \Sigma$ and $(b_1, s_1), (b_0, s_0) \in S_B$ put

$$p'((b_1, s_1) \mid a, (b_0, s_0)) = p_A(b_1, s_1 \mid a, s_0)$$

and

$$\mu(b_2 \mid (b_1, s_1)) = \begin{cases} 1 & \text{if } b_1 = b_2, \\ 0 & \text{otherwise.} \end{cases}$$

Then we have for all $a \in \Sigma$, $b_2 \in \Omega$, and $(b_1, s_1), (b_0, s_0) \in S_B$,

$$p_B(b_2, (b_1, s_1) \mid a, (b_0, s_0)) = \mu(b_2 \mid (b_1, s_1)) \cdot p'((b_1, s_1) \mid a, (b_0, s_0)).$$

It is clear that B is a stochastic Moore automaton. In particular, μ is a mapping, i.e., the successor state determines uniquely the output symbol.

Second, define the mapping $\phi : S_B \to S_A$ by $\phi((b, s)) = s$ for all $(b, s) \in S_B$. Then

$$p_B(b_2, \phi^{-1}\phi(b_1, s_1) \mid a, (b_0, s_0))$$
$$= \sum_{b \in \Omega} p_B(b_2, (b, s_1) \mid a, (b_0, s_0))$$
$$= p_B(b_2, (b_2, s_1) \mid a, (b_0, s_0))$$
$$= p_A(b_2, s_1 \mid a, s_0)$$
$$= p_A(b_2, \phi(b_1, s_1) \mid a, \phi(b_0, s_0)),$$

where $\phi^{-1}\phi(b_1, s_1) = \{(b, s) \in S_B \mid \phi(b, s) = \phi(b_1, s_1)\} = \{(b, s_1) \in S_B \mid b \in \Omega\}$. It follows that the mapping ϕ is an S-epimorphism. In this way, the stochastic automata A and B are S-equivalent. $\quad\square$

4 Cascade Products

The cascade product of deterministic finite-state automata plays an important role in Krohn-Rhodes theory [22, 15, 13, 16]. In this section, a generalization of the cascade product for stochastic automata will be introduced. Note that a related notion of loop-free decomposition of generalized finite-state semiautomata has been studied [23, 24].

The *cascade product* of two stochastic automata $A = (S_A, \Sigma, \Omega_A, p_A)$ and $B = (S_B, \Sigma_B, \Omega_B, p_B)$ with $\Sigma_B = \Sigma \times S_A$ is the stochastic automaton $C = (S, \Sigma, \Omega, p)$, written $C = A \ltimes B$, where $S = S_A \times S_B$, $\Omega = \Omega_A \times \Omega_B$, and

$$p((b_1, b_2), (s_1', s_2') \mid a, (s_1, s_2)) =$$
$$p_A(b_1, s_1' \mid a, s_1) \cdot p_B(b_2, s_2' \mid (a, s_1), s_2) \quad (14)$$

for all $a \in \Sigma$, $(s_1, s_2), (s_1', s_2') \in S$ and $(b_1, b_2) \in \Omega$.

Note that $p(\cdot, \cdot \mid a, (s_1, s_2))$ is a conditional probability distribution, since for all $a \in \Sigma$, $(s_1, s_2), (s_1', s_2') \in S$ and $(b_1, b_2) \in \Omega$,

$$\sum_{b_1, b_2} \sum_{s_1', s_2'} p((b_1, b_2), (s_1', s_2') \mid a, (s_1, s_2))$$
$$= \sum_{b_1, s_1'} p_A(b_1, s_1' \mid a, s_1) \sum_{b_2, s_2'} p_B(b_2, s_2' \mid (a, s_1), s_2)$$
$$= 1.$$

The definition of cascade product can be generalized by introducing a mapping $\omega : \Sigma \times S_A \to \Sigma_B$ such that

$$p((b_1, b_2), (s_1', s_2') \mid a, (s_1, s_2)) =$$
$$p_A(b_1, s_1' \mid a, s_1) \cdot p_B(b_2, s_2' \mid \omega(a, s_1), s_2) \quad (15)$$

for all $a \in \Sigma$, $(s_1, s_2), (s_1', s_2') \in S$ and $(b_1, b_2) \in \Omega$. Then the cascade product is written as $C = A \ltimes_\omega B$.

Proposition 2. *Given stochastic Moore automata* $A = (S_A, \Sigma, \Omega_A, p_A', \mu_A)$ *and* $B = (S_B, \Sigma \times S_A, \Omega_B, p_B', \mu_B)$. *The cascade product* $C = A \ltimes B$ *is a stochastic Moore automaton*

$$C = (S_A \times S_B, \Sigma, \Omega_A \times \Omega_B, p_C', \mu_C)$$

with transition probabilities

$$p_C((b_1, b_2), (s_1', s_2') \mid a, (s_1, s_2)) = \quad (16)$$
$$\mu_C((b_1, b_2) \mid (s_1', s_2')) \cdot p_C'((s_1', s_2') \mid a, (s_1, s_2)),$$

where

$$\mu_C((b_1, b_2) \mid (s_1', s_2')) = \mu_A(b_1 \mid s_1') \cdot \mu_B(b_2 \mid s_2') \,(17)$$
and

$$p_C'((s_1', s_2') \mid a, (s_1, s_2)) =$$
$$p_A'(s_1' \mid a, s_1) \cdot p_B'(s_2' \mid (a, s_1), s_2) \quad (18)$$

for all $a \in \Sigma$ *and* $(s_1, s_2), (s_1', s_2') \in S_C$, *and* $(b_1, b_2) \in \Omega_C$.

Proof. The cascade product $C = A \ltimes B$ has the transition probabilities

$$p_C((b_1, b_2), (s_1', s_2') \mid a, (s_1, s_2)) =$$
$$p_A(b_1, s_1' \mid a, s_1) \cdot p_B(b_2, s_2' \mid (a, s_1), s_2),$$

where $a \in \Sigma$, $(s_1, s_2), (s_1', s_2') \in S_C$, and $(b_1, b_2) \in \Omega_C$. Since A and B are stochastic Moore automata,

$$p_A(b_1, s_1' \mid a, s_1) = \mu_A(b_1 \mid s_1') \cdot p_A'(s_1' \mid a, s_1)$$

and

$$p_B(b_2, s_2' \mid (a, s_1), s_2) =$$
$$\mu_B(b_2 \mid s_2') \cdot p_B'(s_2' \mid (a, s_1), s_2).$$

In this way, eqns. (16)-(18) are established.

Moreover, for each pair $(s_1, s_2) \in S_C$, $\mu_C(\cdot \mid (s_1, s_2))$ is a conditional probability distribution, since for all $s_1, s_2 \in S$,

$$\sum_{b_1, b_2} \mu_C((b_1, b_2) \mid (s_1, s_2))$$
$$= \sum_{b_1} \mu_A(b_1 \mid s_1) \cdot \sum_{b_2} \mu_B(b_2 \mid s_2)$$
$$= 1.$$

Furthermore, for all $(s_1, s_2) \in S_C$ and $a \in \Sigma$, $p_C'(\cdot \mid a, (s_1, s_2))$ is a conditional probability distribution, since for all $a \in \Sigma$ and $s_1, s_2 \in S$,

$$\sum_{s_1', s_2'} p_C'((s_1', s_2') \mid a, (s_1, s_2))$$
$$= \sum_{s_1'} p_A'(s_1' \mid a, s_1) \cdot \sum_{s_2'} p_B'(s_2' \mid (a, s_1), s_2)$$
$$= 1.$$

Finally, $p_C(\cdot, \cdot \mid a, (s_1, s_2))$ is a conditional probability distribution, since for all $a \in \Sigma$ and $s_1, s_2 \in S$,

$$\sum_{b_1, b_2} \sum_{s'_1, s'_2} p_C((b_1, b_2), (s'_1, s'_2) \mid a, (s_1, s_2))$$

$$= \sum_{b_2, s'_2} \sum_{b_1, s'_1} p_A(b_1, s'_1 \mid a, s_1) p_B(b_2, s'_2 \mid (a, s_1), s_2)$$

$$= 1.$$

$\square$

Example 3. *Consider the stochastic Moore automaton*

$$A = (\{s_1, s_2\}, \{a\}, \{b, c\}, p'_A, \mu_A)$$

with probabilities specified by

	$\mu_A(b \mid \cdot)$	$\mu_A(c \mid \cdot)$	$p'_A(s_1 \mid a, \cdot)$	$p'_A(s_2 \mid a, \cdot)$
s_1	1	0	$\frac{1}{3}$	$\frac{2}{3}$
s_2	0	1	$\frac{1}{5}$	$\frac{4}{5}$

and the stochastic Moore automaton

$$B = (\{t_1, t_2\}, \{(a, s_1), (a, s_2)\}, \{b, c\}, p'_B, \mu_B)$$

with probabilities defined by

	$\mu_B(b \mid \cdot)$	$\mu_B(c \mid \cdot)$
t_1	$\frac{1}{2}$	$\frac{1}{2}$
t_2	$\frac{1}{3}$	$\frac{2}{3}$

and

	$p'_B(t_1 \mid (a, s_1), \cdot)$	$p'_B(t_2 \mid (a, s_1), \cdot)$
t_1	$\frac{1}{2}$	$\frac{1}{2}$
t_2	$\frac{1}{3}$	$\frac{2}{3}$

	$p'_B(t_1 \mid (a, s_2), \cdot)$	$p'_B(t_2 \mid (a, s_2), \cdot)$
t_1	$\frac{1}{2}$	$\frac{1}{2}$
t_2	0	1

The cascade product $C = A \ltimes B$ is the stochastic Moore automaton

$$C = (\{s_1, s_2\} \times \{t_1, t_2\}, \{a\}, \{b, c\}^2, p'_C, \mu_C)$$

with probabilities given by

	$\mu_C((b, b) \mid \cdot)$	$\mu_C((b, c) \mid \cdot)$
(s_1, t_1)	$\frac{1}{2}$	$\frac{1}{2}$
(s_1, t_2)	$\frac{1}{3}$	$\frac{2}{3}$
(s_2, t_1)	0	0
(s_2, t_2)	0	0

	$\mu_C((c, b) \mid \cdot)$	$\mu_C((c, c) \mid \cdot)$
(s_1, t_1)	0	0
(s_1, t_2)	0	0
(s_2, t_1)	$\frac{1}{2}$	$\frac{1}{2}$
(s_2, t_2)	$\frac{1}{3}$	$\frac{2}{3}$

and

	$p'_C((s_1, t_1) \mid a, \cdot)$	$p'_C((s_1, t_2) \mid a, \cdot)$
(s_1, t_1)	$\frac{1}{6}$	$\frac{1}{6}$
(s_1, t_2)	$\frac{2}{9}$	$\frac{2}{9}$
(s_2, t_1)	$\frac{1}{10}$	$\frac{1}{10}$
(s_2, t_2)	0	$\frac{1}{5}$

	$p'_C((s_2, t_1) \mid a, \cdot)$	$p'_C((s_2, t_2) \mid a, \cdot)$
(s_1, t_1)	$\frac{1}{3}$	$\frac{1}{3}$
(s_1, t_2)	$\frac{4}{9}$	$\frac{4}{9}$
(s_2, t_1)	$\frac{2}{5}$	$\frac{2}{5}$
(s_2, t_2)	0	$\frac{4}{5}$

$\diamondsuit$

A special case of the cascade product is the *direct composition* of two stochastic automata $A = (S_A, \Sigma_A, \Omega_A, p_A)$ and $B = (S_B, \Sigma_B, \Omega_B, p_B)$, which is given by the cascade product $C = (S, \Sigma, \Omega, p)$, where $S = S_A \times S_B$, $\Sigma = \Sigma_A \times \Sigma_B$, $\Omega = \Omega_A \times \Omega_B$, and the conditional probabilities

$$p((b_1, b_2), (s'_1, s'_2) \mid a, (s_1, s_2))$$
$$= p_A(b_1, s'_1 \mid a, s_1) \cdot p_B(b_2, s'_2 \mid a, s_2) \quad (19)$$

for all $a \in \Sigma$, $(s_1, s_2), (s'_1, s'_2) \in S$ and $(b_1, b_2) \in \Omega$.

A *partition* R of a state set S is a collection $R = \{R_i \mid i \in T\}$ of non-empty subsets R_i of S such that each element of S belongs to exactly one subset. The elements of a partition are called *blocks*. For instance, partitions of the set $\{1, \ldots, 5\}$ are $\{\{1, 2\}, \{3, 4\}, \{5\}\}$ and $\{\{1, 2, 3\}, \{4\}, \{5\}\}$.

The set of partitions of a non-empty set S can be *partially ordered* such that for any partitions R_1, R_2 of S, $R_1 \leq R_2$ if each block of R_2 is the union of one or more blocks of R_1. The smallest partition is $\underline{0} = \{\{s\} \mid s \in S\}$ and the largest partition is $\underline{1} = \{S\}$. Thus for each partition R of S, we have $\underline{0} \leq R \leq \underline{1}$. More specifically, the partitions of a non-empty set S form a *lattice*, i.e., any two partitions have a least upper bound and a greatest lower bound denoted by inf and sup, respectively. For instance, $\underline{0} \leq \{\{1, 2\}, \{3\}, \{4\}, \{5\}\} \leq \{\{1, 2\}, \{3, 4\}, \{5\}\} \leq \{\{1, 2, 5\}, \{3, 4\}\} \leq \underline{1}$.

In view of a stochastic Moore automaton A, a partition Q of the state set S_A has the *substitution property* for A if for each input $a \in \Sigma$, all blocks $Q_i, Q_j \in Q$, and states $s_1, s_2 \in Q_i$, marginal probabilities fulfill

$$\sum_{s' \in Q_j} p'_A(s' \mid a, s_1) = \sum_{s' \in Q_j} p'_A(s' \mid a, s_2). \quad (20)$$

Two partitions Q and R of the state set S_A are *mutually independent* if for all $a \in \Sigma$, $s \in S_A$ and each

pair of blocks Q_i and R_j,

$$\sum_{s' \in Q_i \cap R_j} p'_A(s' \mid a, s)$$
$$= \sum_{s' \in Q_i} p'_A(s' \mid a, s) \cdot \sum_{s' \in R_j} p'_A(s' \mid a, s). \quad (21)$$

Theorem 3. *A stochastic Moore automaton $C = (S, \Sigma, \Omega, p, \mu)$ admits a decomposition into a cascade product $C = A \ltimes B$ of stochastic Moore automata A and B if there are partitions Q and R of the state set S such that Q has the substitution property for C, Q and R are mutually independent, Q and R are output compatible (see proof), and $\inf(Q, R) = \underline{0}$.*

Proof. Let $Q = \{Q_1, \ldots, Q_m\}$ and $R = \{R_1, \ldots, R_n\}$. In view of the condition $\inf(Q, R) = \underline{0}$, the cartesian product $Q \times R = \{(Q_i, R_j) \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ corresponds one-to-one with the state set S, where the pair (Q_j, R_i) with $Q_i \cap R_j = \{s_{ij}\}$ is associated with the state s_{ij}. Thus $S_C = \{s_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$.

Define the stochastic Moore automaton $A = (Q, \Sigma, \Omega_Q, p, \mu)$, where

$$p'_A(Q_j \mid a, Q_i) = \sum_{s' \in Q_j} p'_C(s' \mid a, s), \quad s \in Q_i. \quad (22)$$

By the substitution property, the conditional distribution p'_A is well-defined.

Define the stochastic Moore automaton $B = (R, \Sigma \times Q, \Omega_R, p, \mu)$, where

$$p'_B(R_l \mid (a, Q_i), R_k) = \sum_{s' \in R_l} p'_C(s' \mid a, s), \quad (23)$$

$$Q_i \cap R_k = \{s\}.$$

Then by the mutual independence of Q and R,

$$p'_C((Q_j, R_l) \mid a, (Q_i, R_k))$$
$$= p'_C(s_{jl} \mid a, s_{ik}) \quad (24)$$
$$= \sum_{s' \in Q_j} p'_C(s' \mid a, s_{ik}) \cdot \sum_{s' \in R_l} p'_C(s' \mid a, s_{ik})$$
$$= p'_A(Q_j \mid a, Q_i) \cdot p'_B(R_l \mid (a, Q_i), R_k).$$

Furthermore, the partitions Q and R of the state set S are *output compatible*, i.e., the output set of C can be written as $\Omega = \Omega_Q \times \Omega_R$ such that for each $b_1 \in \Omega_Q, b_2 \in \Omega_R$ and $Q_i \in Q, R_j \in R$,

$$\mu_C((b_1, b_2) \mid (Q_i, R_j)) =$$
$$\mu_A(b_1 \mid Q_i) \cdot \mu_B(b_2 \mid R_j), \quad (25)$$

where μ_A and μ_B are conditional probabilities on Q and R, respectively. Hence, C is the cascade product of A and B. $\qquad\square$

Example 4. *Consider the stochastic Moore automaton*

$$C = (\{s_1, \ldots, s_4\}, \{a\}, \{(b_1, c_1),$$
$$(b_1, c_2), (b_2, c_1), (b_2, c_2)\}, p', \mu)$$

with probabilities given by

	$\mu_C((b_1, c_1) \mid \cdot)$	$\mu_C((b_1, c_2) \mid \cdot)$
s_1	$\frac{1}{4}$	$\frac{1}{4}$
s_2	$\frac{1}{8}$	$\frac{3}{8}$
s_3	$\frac{1}{6}$	$\frac{1}{6}$
s_4	$\frac{1}{12}$	$\frac{1}{4}$

	$\mu_C((b_2, c_1) \mid \cdot)$	$\mu_C((b_2, c_2) \mid \cdot)$
s_1	$\frac{1}{4}$	$\frac{1}{4}$
s_2	$\frac{1}{8}$	$\frac{3}{8}$
s_3	$\frac{1}{3}$	$\frac{1}{3}$
s_4	$\frac{1}{6}$	$\frac{1}{2}$

and

	$p'_C(s_1 \mid a, \cdot)$	$p'_C(s_2 \mid a, \cdot)$
s_1	$\frac{1}{8}$	$\frac{3}{8}$
s_2	$\frac{1}{8}$	$\frac{3}{8}$
s_3	$\frac{1}{5}$	$\frac{1}{5}$
s_4	$\frac{1}{10}$	$\frac{3}{10}$

	$p'_C(s_3 \mid a, \cdot)$	$p'_C(s_4 \mid a, \cdot)$
s_1	$\frac{1}{8}$	$\frac{3}{8}$
s_2	$\frac{1}{8}$	$\frac{3}{8}$
s_3	$\frac{1}{5}$	$\frac{2}{5}$
s_4	$\frac{2}{5}$	$\frac{1}{5}$

The partitions

$$Q = \{Q_1 = \{s_1, s_2\}, Q_2 = \{s_3, s_4\}\}$$
$$\text{and } R = \{R_1 = \{s_1, s_3\}, R_2 = \{s_2, s_4\}\}$$

satisfy $\inf(Q, R) = \underline{0}$.

Define the stochastic Moore automaton

$$A = (\{Q_1, Q_2\}, \{a\}, \{b_1, b_2\}, p', \mu)$$

with probabilities given by

	$\mu_A(b_1 \mid \cdot)$	$\mu_A(b_2 \mid \cdot)$
Q_1	$\frac{1}{2}$	$\frac{1}{2}$
Q_2	$\frac{1}{3}$	$\frac{2}{3}$

and

	$p'_A(Q_1 \mid a, \cdot)$	$p'_A(Q_2 \mid a, \cdot)$
Q_1	$\frac{1}{2}$	$\frac{1}{2}$
Q_2	$\frac{1}{4}$	$\frac{3}{4}$

Moreover, define the stochastic Moore automaton

$$B = (\{R_1, R_2\}, \{(a, Q_1), (a, Q_2)\}, \{c_1, c_2\}, p', \mu)$$

with probabilities given by

	$\mu_B(c_1 \mid \cdot)$	$\mu_B(c_2 \mid \cdot)$
R_1	$\frac{1}{2}$	$\frac{1}{2}$
R_2	$\frac{1}{4}$	$\frac{3}{4}$

and

	$p'_B(R_1 \mid (a, Q_1), \cdot)$	$p'_B(R_1 \mid (a, Q_2), \cdot)$
R_1	$\frac{1}{6}$	$\frac{1}{6}$
R_2	$\frac{1}{6}$	$\frac{1}{6}$

	$p'_B(R_2 \mid (a, Q_1), \cdot)$	$p'_B(R_2 \mid (a, Q_2), \cdot)$
R_1	$\frac{1}{3}$	$\frac{1}{3}$
R_2	$\frac{1}{3}$	$\frac{1}{3}$

$\diamond$

5 Conclusion

The classes of stochastic Moore automata and stochastic automata have the same functionality, but stochastic Moore automata have the advantage that they are easier to access than their pure stochastic counterparts. Cascade decomposition of automata plays an important role for their analysis. This paper has shown that stochastic Moore automata can be decomposed into a cascade product of stochastic Moore automata under mild conditions. Iterated cascade decomposition might lead to a loop-free partitioning of stochastic Moore automata and in this way might contribute to the exploration of complex systems.

References:

[1] C. E. Shannon: The mathematical theory of communication, *Bell System Technical Journal*, Vol.5, No.1, 1948, pp. 379-423. doi: 10.1002/j.1538-7305.1948.tb01338.x

[2] J. von Neumann: Probabilistic logic and the synthesis of reliable organisms from unreliable components, in: Automata Studies, C. Shannon and J. McCarthy (eds), *Annals of Mathematical Studies*, Vol.34, Princeton Univ. Press, Princeton, NJ, 1956. doi: 10.1515/9781400882618-003

[3] M. O. Rabin, D. Scott: Finite automata and their decision problems, *IBM Journal Research Development*, Vol.3, No.3, 1959, pp. 114-125. doi: 10.1147/rd.32.0114

[4] J. W. Carlyle: Reduced forms for stochastic sequential machines, *Journal Mathematical Analysis and Applications*, Vol.7, No.2, 1963, pp. 167-165. doi: 10.1016/0022-247X(63)90045-3

[5] P. H. Starke: Stochastische Ereignisse und Wortmengen, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, Vol.12, 1966, pp. 61-68. doi: 10.1002/malq.19660120108

[6] Felix X.-F. Ye, F. Wang, Hong Qian: Stochastic dynamics: Markov chains and random transformations, *Discrete and Continuous Dynamical Systems - Series B*, Vol.21, No.7, 2016, pp. 2337-2361.

[7] M. O. Rabin: Probabilistic automata, *Information and Control*, Vol.6, No.3, 1963, pp. 230-245. doi: 10.1016/S0019-9958(63)90290-0

[8] P. Turakainen: Generalized automata and stochastic languages, *Proc. Amer. Math. Soc.*, Vol.21, 1969, pp. 303-309. doi: 10.2307/2036989

[9] M. N. Cakir, M. Saleemi, K.-H. Zimmermann: On the theory of stochastic automata, Preprint, TU Hamburg, arxiv:2103.14423, 2021.

[10] V. Claus: *Stochastische Automaten*, Teubner, Stuttgart, 1971.

[11] A. Salomaa: *Theory of Automata*, Pergamon Press, Oxford (1969).

[12] M. N. Cakir, K.-H. Zimmermann: On stochastic automata over monoids, Preprint, TU Hamburg, arxiv:2002.01214, 2020.

[13] V. Diekert, M. Kufleitner, B. Steinberg: The Krohn-Rhodes theorem and local divisors, *Fundam. Inform.*, Vol.116, 2012, pp. 65-77. doi: 10.1016/s0304-3975(99)00315-1

[14] S. Mihov, K.U. Schulz: *Finite-State Techniques*, Cambridge Univ. Press, New York, 2019.

[15] A. Ginzburg: *Algebraic Theory of Automata*, Academic Press, New York, 1968.

[16] K.-H. Zimmermann: On Krohn-Rhodes theory for semiautomata, Preprint, TU Hamburg, arXiv:2103-14423, 2020.

[17] A. Akhavi, I. Klimann, S. Lombardy, J. Mairesse, M. Piantin: On the finiteness problem for automaton (semi)groups, *Int. J. Algebra Comp.*, Vol.22, No.6, 2012.

[18] J.H. Conway: *Regular Algebra and Finite Machines*, Dover Publications, London, 2012.

[19] E.A. Lee, S.A. Seshia: *Introduction to Embedded Systems*, UC Berkely, Lulu.com, 2013.

[20] P. H. Starke: Theory of stochastic automata, *Kybernetika*, Vol.6, No.2, 1966, pp. 475-482.

[21] K.-H. Zimmermann: Inference and learning in stochastic automata, *Int. J. Pure and Applied Mathematics*, Vol.115, No.3, 2017, pp. 621-639.

[22] N. Ghadbane: Transformation semigroups and state machines, *Applied Mathematics and Computational Intelligence*, Vol.8, 2019, pp. 9-16.

[23] R. G. Bukharaev: *Theorie der stochastischen Automaten*, Teubner, Stuttgart, 1995.

[24] M. N. Cakir, M. Saleemi, K.-H. Zimmermann: On the decomposition of generalized semiautomata, *WSEAS Trans. Inform. Sci. Appl.*, Vol.18, 2021, pp. 34-38. doi: 10.37394/23209.2021