International Journal of Control
Vol. 80, No. 5, May 2007, 789–799

Generalized language measure families of probabilistic finite state systems

I. CHATTOPADHYAY and A. RAY*

Mechanical Engineering Department, Pennsylvania State University, University Park, PA 16802, USA

(*Received 11 August 2006; in final form 28 December 2006*)

The signed real measure of regular languages has been introduced and validated in recent literature for quantitative analysis of discrete-event systems. This paper reports generalizations of the language measure, which can serve as performance indices for synthesis of optimal discrete-event supervisory decision and control laws. These generalizations eliminate a user-selectable parameter in the original concept of language measure. The concepts are illustrated with simple examples.

1. Introduction

In the discrete-event setting, a finite-state automaton (FSA) model of a physical plant is a generator of its regular language, whose behaviour is constrained by the supervisor (or controller) to meet a given specification. A signed real measure of regular languages has been reported in Ray (2005) and Ray *et al.* (2005) to provide a mathematical framework for quantitative comparison of controlled sublanguages. In this work, each transition is assigned a cost, similar to its probability measure that can be quantitatively evaluated from physical experimentation or extensive simulation on a test bed. Each state of the FSA model is assigned a signed real weight whose upper and lower bounds are normalized to 1 and -1, respectively. The measure of a given event trace is obtained as the product of the cost of transitions and the (normalized) weight of the terminating state. The sum of the measures of all traces yield the language measure.

Optimal control of finite state automata has been recently reported Ray *et al.* (2004, 2005) based on the total ordering induced by the language measure as augmentation to the supervisory control theory of Ramadge and Wonham (1987). This work consolidates the theory and applications of optimal supervisory control of regular languages, where the performance index is obtained by combining a real signed measure of the supervised plant language with the cost of disabled event(s). Starting with the (regular) language of an unsupervised plant automaton, the optimal control policy makes a trade-off between the measure of the supervised sublanguage and the associated event disabling cost to achieve the best performance. Like any other optimization procedure, it is possible to choose different performance indices to arrive at different optimal policies for discrete event supervisory control. It is recognized that optimal control of discrete-event systems can be achieved with a cost function that may not qualify as a measure (e.g., Sengupta and Lafortune (1998)). Nevertheless, usage of a language measure as the cost function facilitates precise comparative evaluation of different supervisors so that the appropriate control policy(ies) can be conclusively identified.

From the above perspectives, this paper presents generalizations of the language measure (Ray 2005), each generalization being a formal measure in its own right and having physical implications that are relevant to synthesis of discrete-event supervisory control policies. These generalizations are achieved through a new concept of trace measure that is characterized by both initiating and terminating states as well as the length of the trace and the choice of a vector

*Corresponding author. Eamil: axr2@psu.edu

International Journal of Control
ISSN 0020–7179 print/ISSN 1366–5820 online © 2007 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/00207170601188784

norm (Naylor and Sell 1982). The concept of generalization can be viewed as renormalization (Chattopadhyay and Ray 2006) of the (normalized) language measure (Ray 2005).

The paper is organized in six sections including the present one. Section 2 briefly reviews background concepts on language measure. Section 3 derives measures related to the stationary state probability vector of the finite-state automaton. Section 4 introduces the notion of shaped measures, which allows assignment of selective length-based importance to different traces in the generated language. It is further shown that measures introduced in § 3 can be obtained as limits of sequences of shaped measures. Section 5 presents an example of optimal automaton configurations. Section 6 concludes the paper along with recommendations for future research.

2. Brief review of language measure

This section briefly reviews the concept of signed real measure of regular languages Ray (2005) and Ray *et al.* (2005). Let $G_i \equiv \langle Q, \Sigma, \delta, q_i, Q_m \rangle$ be a trim (i.e., accessible and co-accessible) deterministic finite-state automaton (DFSA) model (Ramadge and Wonham 1987) that represents the discrete-event dynamics of a physical plant, where $Q = \{q_k : k \in I_Q\}$ is the set of states and $I_Q \equiv \{1, 2, \ldots, n\}$ is the index set of states; the automaton starts with the initial state q_i; the alphabet of events is $\Sigma = \{\sigma_k : k \in I_\Sigma\}$, and $I_\Sigma \equiv \{1, 2, \ldots, \ell\}$ is the index set of events; δ: $Q \times \Sigma \to Q$ is the (possibly partial) function of state transitions; and $Q_m \equiv \{q_{m_1}, q_{m_2}, \ldots, q_{m_r}\} \subseteq Q$ is the set of marked (i.e., accepted) states with $q_{m_k} = q_j$ for some $j \in I_Q$.

Let $\Sigma^\star$ be the Kleene closure of Σ, i.e., the set of all finite-length traces made of the events belonging to Σ as well as the empty trace ϵ that is viewed as the identity of the monoid $\Sigma^\star$ under the operation of trace concatenation, i.e., $\epsilon s = s = s\epsilon$. The extension $\delta^\star$: $Q \times \Sigma^\star \to Q$ is defined recursively in the usual sense (Hppcroft *et al.* 2001).

Definition 1: The language $L(G_i)$ generated by a *DFSA* G initialized at the state $q_i \in Q$ is defined as $L(G_i) = \{s \in \Sigma^\star \mid \delta^\star(q_i, s) \in Q\}$.

Definition 2: The language $L_m(G_i)$ marked by a *DFSA* G_i initialized at the state $q_i \in Q$ is defined as $L_m(G_i) = \{s \in \Sigma^\star \mid \delta^\star(q_i, s) \in Q_m\}$.

The language $L(G_i)$ is partitioned into non-marked and marked languages, $L^o(G_i) \equiv L(G_i) - L_m(G_i)$ and $L_m(G_i)$, consisting of event traces that, starting from $q_i \in Q$, terminate at one of the non-marked states in $Q - Q_m$ and one of the marked states in Q_m, respectively. The set Q_m is further partitioned into Q_m^+ and Q_m^-, where Q_m^+ contains all good marked states that are desired to be terminated on and Q_m^- contains all bad marked states that one may not want to terminate on, although it may not always be possible to avoid the bad states while attempting to reach the good states. Accordingly, the marked language $L_m(G_i)$ is further partitioned into $L_m^+(G_i)$ and $L_m^-(G_i)$ consisting of good and bad traces that, starting from q_i, terminate on Q_m^+ and Q_m^-, respectively. Thus, the language $L(G_i)$ is decomposed into null, i.e., $L^o(G_i)$, positive, i.e., $L_m^+(G_i)$, and negative, i.e., $L_m^-(G_i)$ sublanguages. A signed real measure μ: $2^{L(G_i)} \to \mathbb{R} \equiv (-\infty, \infty)$ is constructed for quantitative evaluation of every event trace $s \in L(G_i)$.

Definition 3: The language of all traces that, starting at a state $q_i \in Q$, terminates on a state $q_j \in Q$, is denoted as $L(q_i, q_j)$. That is, $L(q_i, q_j) \equiv \{s \in L(G_i) : \delta^\star(q_i, s) = q_j\}$.

Definition 4: The terminating characteristic function that assigns a normalized signed real weight to state-partitioned sublanguages $L(q_i, q_j)$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$ is defined as χ: $Q \to [-1,\ 1]$ such that

$$\chi_j \in \begin{cases} [-1,\ 0) & \text{if } q_j \in Q_m^- \\ \{0\} & \text{if } q_j \notin Q_m \\ (0,\ 1] & \text{if } q_j \in Q_m^+. \end{cases} \tag{1}$$

Definition 5: The event cost is conditioned on a *DFSA* state at which the event is generated, and is defined as $\tilde{\pi}: L(G_i)) \times Q \to [0,1]$ such that $\forall q_j \in Q, \forall \sigma_k \in \Sigma$, $\forall s \in L(G_i))$

(1) $\tilde{\pi}[\sigma_k, q_j] \equiv \tilde{\pi}_{jk} \in [0,1)$; $\sum_k \tilde{\pi}_{jk} < 1$;
(2) $\tilde{\pi}[\sigma, q_j] = 0$ if $\delta(q_j, \sigma)$ is undefined; $\tilde{\pi}[\epsilon, q_j] = 1$;
(3) $\tilde{\pi}[\sigma_k s, q_j] = \tilde{\pi}[\sigma_k, q_j]\ \tilde{\pi}[s, \delta(q_j, \sigma_k)]$.

The event cost matrix is defined as $\widetilde{\mathbf{\Pi}}_{ij} = \tilde{\pi}_{ij}$ with $i \in \{1, \ldots, n\}$ and $j \in \{1, \ldots, m\}$ where the automaton has n states and cardinality of the event alphabet Σ is m.

An application of the induction principle to part (3) in Definition 5 shows $\tilde{\pi}[st, q_j] = \tilde{\pi}[s, q_j]\tilde{\pi}[t, \delta^\star(q_j, s)]$. The condition $\sum_k \tilde{\pi}_{jk} < 1$ provides a sufficient condition for the existence of the real signed measure (Ray 2005). Next a measure of sublanguages of the plant language $L(G_i)$ is formulated in terms of the signed characteristic function χ and the non-negative event cost $\tilde{\pi}$.

Definition 6: The state transition cost, π: $Q \times Q \to [0,1)$, of the *DFSA* G_i is defined as follows: $\forall q_i, q_j \in Q$,

$$\pi_{ij} = \begin{cases} \sum_{\sigma \in \Sigma} \tilde{\pi}[\sigma, q_i], & \text{if } \delta(q_i, \sigma) = q_j \\ 0 & \text{if}\{\delta(q_i, \sigma) = q_j\} = \emptyset. \end{cases} \tag{2}$$

Consequently, the $n \times n$ state transition cost $\mathbf{\Pi}$-matrix is defined as $\mathbf{\Pi}_{ij} = \pi_{ij}$ with $i, j \in \{1, \ldots, n\}$ where the number of states in the automaton is n.

Although the preceding analysis reported in Ray (2005) and Ray *et al.* (2005) was intended for non-probabilistic regular languages, the event costs can be interpreted as conditional probabilities of event occurrence. A brief discussion on the physical interpretation of the event costs is given in Ray (2005) to explain this issue. Furthermore, an element π_{jk} of the Π-matrix is conceptually similar to the state transition probability of a Markov chain or a semi-Markov chain with the exception that the equality condition $\sum_k \pi_{jk} = 1$ is not satisfied. Specifically, the inequality $\sum_k \pi_{jk} < 1,\ j = 1, 2, \ldots, n$ provides a sufficient condition for the language measure to be finite. This implies that the preceding analysis is applicable to the case of terminating probabilistic languages (Garg 1992a, b) that have a non-zero probability of termination (arising from either intentional design or unmodelled dynamics of the plant automaton) at each state. If the probability of termination at each state, or equivalently the probability of transition to the (deadlock) dump state from each of the other states $q_i \in Q$, is set identically equal to $\theta \in (0, 1)$, then the $\widetilde{\mathbf{\Pi}}$-matrix and the $\mathbf{\Pi}$-matrix can be θ-parameterized as follows (Chattopadhyay and Ray 2006):

$$\widetilde{\mathbf{\Pi}}(\theta) \equiv (1-\theta)\widetilde{\mathbf{P}} \quad \text{and} \quad \mathbf{\Pi}(\theta) \equiv (1-\theta)\mathbf{P}, \tag{3}$$

where $\widetilde{\mathbf{P}}$ is the event matrix (also known as the morph matrix), which is derived from experimental data or simulation data (Ray 2005) and the resulting stochastic state transition matrix $\mathbf{P}$ is obtained from $\widetilde{\mathbf{P}}$ in a way similar to equation (2). Since $\mathbf{P}$ is a stochastic matrix (i.e., $\sum_j \mathbf{P}_{ij} = 1\ \forall i \in \{1, \ldots, n\}$), the row sums $\sum_j \pi_{ij} = (1-\theta) < 1,\ j = 1, 2, \ldots, n$ (see Definition 6) make $\mathbf{\Pi}$ a contraction operator with the magnitude of each of its eigenvalues being less than or equal to $(1-\theta)$; consequently, $[\mathbb{I} - \mathbf{\Pi}]$ becomes invertible (Ray 2005).

In the sequel, the preceding measure construction is generalized and the notion of language measure is extended to non-terminating models by first assuming a uniform non-zero probability of termination θ at each state and then computing the limit as $\theta \to 0^+$, i.e., the probability of termination approaching zero. The resulting θ-parameterized model coincides with the desired non-terminating model in the limit (Chattopadhyay and Ray 2006).

Definition 7: The θ-parameterized measure of the language $L(q_i, q_j)$ is defined in terms of its traces (see Definitions 3, 4 and 5) as

$$\mu^\theta(\{s\}) \equiv \tilde{\pi}(s, q_i)\chi_j, \quad \forall s \in L(q_i, q_j) \tag{4}$$

$$\mu^\theta\big(L(q_i, q_j)\big) \equiv \sum_{s \in L(q_i, q_j)} \mu^\theta(\{s\}). \tag{5}$$

Then, the measure of the language $L(G_i)$ of a *DFSA* G_i, initialized at the state $q_i \in Q$, is defined as

$$\mu^\theta(L(G_i)) = \sum\nolimits_j \mu^\theta\big(L(q_i, q_j)\big) \tag{6}$$

It is shown in Ray (2005) that the measure $\mu_i^\theta \equiv \mu^\theta(L(G_i))$ can be expressed as: $\mu_i^\theta = \sum_j \pi_{ij}\mu_j^\theta + \chi_i$. In vector notation, the θ-parameterized language measure vector is expressed by making use of equation (3) as

$$\boldsymbol{\mu}^\theta = [\mathbb{I} - (1-\theta)\mathbf{P}]^{-1}\boldsymbol{\chi}, \tag{7}$$

where the measure vector $\boldsymbol{\mu}^\theta \equiv [\mu_1^\theta\ \mu_2^\theta\ \cdots\ \mu_n^\theta]^T$ and the terminating characteristic vector $\boldsymbol{\chi} \equiv [\chi_1\ \chi_2\ \cdots\ \chi_n]^T$. Note that $\lim_{\theta \to 0^+} \boldsymbol{\mu}^\theta$ of the normalized language measure does not exist. This problem has been circumvented via renormalization (Chattopadhyay and Ray 2006) as explained below.

The regular language $L(G_i)$ is a sublanguage of the Kleene closure $\Sigma^\star$ of the alphabet Σ, for which the automaton states can be merged into a single state. Then, $\mathbf{P}$ degenerates to the 1×1 identity matrix and the terminating characteristic vector $\boldsymbol{\chi}$ becomes one-dimensional and can be assigned as $\boldsymbol{\chi} = 1$ by normalization. Consequently, the measure vector $\boldsymbol{\mu}^\theta$ in equation (7) degenerates to a scalar measure θ^{-1}. The renormalized measure is obtained from equation (7) after normalization with respect to θ^{-1}.

$$\boldsymbol{\vartheta}^\theta|_1 = \theta\ [\mathbb{I} - (1-\theta)\mathbf{P}]^{-1}\boldsymbol{\chi}. \tag{8}$$

3. Generalization of language measure

This section generalizes the notion of language measure μ^θ (see Definition 7 and equation (7)), which also leads to a renormalized measure ϑ^θ (see equation (8)). This is achieved by redefining the measure of individual traces in terms of an initiating characteristic function ξ: $Q \longmapsto [0, 1]$ that assigns a positive weight to each initiating state q_i and serves as a renormalizing factor (i.e., a multiplicative constant) for the measure of the traces initiating from the respective state. Figure 1 illustrates the relationship among the initiating and terminating characteristics. Different initiating

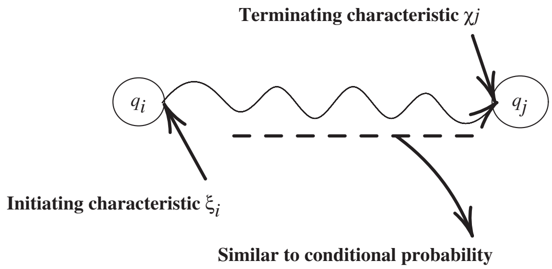

Figure 1. Generalization of langauge measure.

characteristics lead to different renormalized language measures that may have different physical interpretations.

Definition 8: The θ-parameterized generalized measure of a singleton event trace set $\{s\} \subseteq L(q_i, q_j) \subseteq L(G_i)$ in the σ-algebra $2^{L(G_i)}$ is defined as

$$\vartheta^{\theta}(\{s\}) \equiv \xi_i \mu^{\theta}(\{s\}) = \xi(q_i)\tilde{\pi}(s, q_i)\chi_j, \quad \forall s \in L(q_i, q_j). \quad (9)$$

The generalized measure of $L(q_i, q_j)$ is defined as

$$\vartheta^{\theta}\big(L(q_i, q_j)\big) \equiv \sum_{s \in L(q_i, q_j)} \vartheta^{\theta}(\{s\}). \quad (10)$$

The generalized measure of a *DFSA* G_i, initialized at the state $q_i \in Q$, is denoted as $\vartheta_i^{\theta} \equiv \vartheta^{\theta}(L(G_i)) = \sum_j \vartheta^{\theta}(L(q_i, q_j))$.

Now it is ascertained that Definition 8 satisfies the properties of a measure on the defined σ-algebra.

Proposition 1: *The generalized measure $\vartheta^{\theta}: 2^{L(G_i)} \longmapsto \mathbb{R}$ is defined on the measure space $(L(G_i), 2^{L(G_i)}, \vartheta^{\theta})$.*

Proof: It suffices to establish σ-additivity from the following fact. For a fixed $\theta \in (0, 1)$, ϑ_i^{θ} is the product of ξ_i (which is a constant) and μ_i^{θ} which is a signed real measure on the σ-algebra $2^{L(G_i)}$. □

A special family of initiating characteristic functions is considered for the generalized language measure.

Definition 9: Let $\Psi(\theta) \equiv [\mathbb{I} - (1-\theta)\mathbf{P}]^{-1}$. The ℓ_p-family of initiating characteristic functions is defined as

$$\xi_i^p(\theta) = \big|\big|\Psi(\theta)_{i\circ}\big|\big|_p^{-1} \quad \forall p \in [1, \infty], \quad \forall \theta \in (0, 1), \quad (11)$$

where $\| \bullet \|_p$ denotes the ℓ_p-norm of $\bullet$; and ith row of a matrix $\mathbf{M}$ is denoted as $\mathbf{M}_{i\circ}$ and the jth column as $\mathbf{M}_{\circ j}$.

Remark 1: Note that $\lim_{\theta \to 0^+} \xi_i^p(\theta)$ does not exist due to non-invertibility of the operator $[\mathbb{I} - \mathbf{P}]$. However, $(\theta^{-1}\xi_i^p(\theta))|_{\theta=0}$ is well-defined by virtue of norm continuity (Naylor and Sell 1982) and hence $\lim_{\theta \to 0^+} \theta^{-1}\xi_i^p(\theta)$ exists.

Lemma 1: *For $p = 1$, the initiating characteristic*

$$\xi_i^1(\theta) = \theta, \quad \forall i \forall \theta \in (0, 1]. \quad (12)$$

Proof: Let $\boldsymbol{e} = [1 \cdots 1]^T$. Since $\mathbf{P}$ is stochastic, non-negativity of $\Psi(\theta)$ follows from the following expansion:

$$\Psi(\theta) = \sum_{k=0}^{\infty}(1-\theta)^k \mathbf{P}^k \quad \forall \theta \in (0, 1]$$

$$\Longrightarrow \Psi(\theta)\boldsymbol{e} = \sum_{k=0}^{\infty}(1-\theta)^k \mathbf{P}^k \boldsymbol{e} = \sum_{k=0}^{\infty}(1-\theta)^k \boldsymbol{e} = \theta^{-1}\boldsymbol{e}$$

which implies that $\| \Psi(\theta)_{i\circ} \|_1 = \theta^{-1} \ \ \forall i \Longrightarrow \xi_i^1 = \theta$. □

The θ-parameterized generalized measure for $p = 1$ is obtained in the vector notation as

$$\boldsymbol{\vartheta}^{\theta}|_1 = \theta[\mathbb{I} - (1-\theta)\mathbf{P}]^{-1}\boldsymbol{\chi} \quad (13)$$

which is identical to the renormalized measure in equation (8).

In general, the θ-parameterized generalized measure for $p \in [1, \infty]$ is obtained in the matrix notation as

$$\boldsymbol{\vartheta}^{\theta}|_p = \begin{bmatrix} \xi_1^p(\theta) & \cdots & 0 \\ \vdots & \xi_i^p(\theta) & \vdots \\ 0 & \cdots & \xi_n^p(\theta) \end{bmatrix} [\mathbb{I} - (1-\theta)\mathbf{P}]^{-1}\boldsymbol{\chi}. \quad (14)$$

Non-negativity of $\mathbf{P}$ and invertibility of $[I - (1-\theta)\mathbf{P}]$ guarantee that $\|\Psi(\theta)_{i\circ}\|_p \in (0, \infty) \ \forall i$, which implies $\xi_i^p(\theta) \in (0, \infty) \ \forall \ p \in [1, \infty] \ \forall \ \theta \in (0, \infty)$.

3.1 *Limiting values of $\boldsymbol{\vartheta}^{\theta}|_p$ as $\theta \to 0^+$*

This section computes the generalized measures $\boldsymbol{\vartheta}^{\theta}|_p$ as $\theta \to 0^+$, based on the state transition probability matrix $\mathbf{P}$ of a stationary Markov chain with finitely many states. Then, $\mathbf{P}$ is a stochastic matrix. That is, $\mathbf{P}$ is non-negative with each row sum being identically equal to unity (Bapat and Raghavan 1997).

Proposition 2: *For every stochastic matrix* $\mathbf{P}$, *the following limit exists*

$$\lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} \mathbf{P}^j = \mathscr{P}, \quad (15)$$

where $\mathscr{P}$ is a stochastic matrix. Furthermore, $\mathscr{P}$ commutes with $\mathbf{P}$ and is idempotent. That is,

$$\mathbf{P}\mathscr{P} = \mathscr{P}\mathbf{P} = \mathscr{P} = \mathscr{P}^2. \quad (16)$$

Proof: The proof is given in Bapat and Raghavan (1997). □

Since $\mathbf{P}$ is a stochastic matrix, $[\mathbb{I}-\mathbf{P}]e=0$ where $e \equiv [1,1,\ldots,1]^T$. Therefore, $[\mathbb{I}-\mathbf{P}]$ is not invertible for any stochastic matrix $\mathbf{P}$; however, $[\mathbb{I}-(1-\theta)\mathbf{P}]$ is always invertible for $\theta \in (0,1)$. The lemma to the next proposition shows that $[\mathbb{I}-\mathbf{P}+\mathscr{P}]$ is invertible.

Proposition 3: *The matrix* $[\mathbb{I}-\mathbf{P}+\alpha\mathscr{P}]$ *is invertible for all* $\alpha \neq 0$.

Proof: The proof is based on the commutative and idempotent properties of $\mathscr{P}$ in equation (16) and uses the principle of contradiction.

Let $[\mathbb{I}-\mathbf{P}+\alpha\mathscr{P}]$ be non-invertible for an arbitrary $\alpha \neq 0$. Then, there is a vector $\vartheta \neq 0$ such that

$$\begin{aligned}&[\mathbb{I}-\mathbf{P}+\alpha\mathscr{P}]\vartheta = 0\\ &\Rightarrow [\mathbf{P}-\alpha\mathscr{P}]\vartheta = \vartheta \Rightarrow \alpha\mathscr{P}[\mathbf{P}-\alpha\mathscr{P}]\vartheta = \alpha\mathscr{P}\vartheta\\ &\Rightarrow [\alpha\mathscr{P}-\alpha^2\mathscr{P}]\vartheta = \alpha\mathscr{P}\vartheta\\ &\Rightarrow \alpha^2\mathscr{P}\vartheta = 0 \Rightarrow \mathscr{P}\vartheta = 0 \text{ because } \alpha \neq 0.\end{aligned}$$

Hence, $\mathbf{P}\vartheta = \mathbf{P}\vartheta - \alpha\mathscr{P}\vartheta = [\mathbf{P}-\alpha\mathscr{P}]\vartheta = \vartheta$

$$\Rightarrow \mathbf{P}^k\vartheta = \vartheta \quad \forall k \in \mathbb{N}\cup\{0\},$$

which implies

$$\left(\frac{1}{k}\sum_{j=0}^{k-1}P^j\right)\vartheta = \vartheta\ \forall k \Longrightarrow \lim_{k\to\infty}\left(\frac{1}{k}\sum_{j=0}^{k-1}P^j\right)\vartheta = \mathscr{P}\vartheta = \vartheta$$
$$\Longrightarrow \vartheta = 0 \quad \text{because } \mathscr{P}\vartheta = 0.$$

This is a contradiction. □

Lemma 2: *The matrix* $[\mathbb{I}-\mathbf{P}+\mathscr{P}]$ *is invertible.*

Proof: The proof follows by setting $\alpha = 1$ in Proposition 3. □

Proposition 4:

$$[\mathbf{P}-\alpha\mathscr{P}]^k = \mathbf{P}^k - \left[1-(1-\alpha)^k\right]\mathscr{P}, \quad \forall k \in \mathbb{N}\ \forall\alpha \neq 0. \tag{17}$$

Proof: The above identity is valid for $k=0$ and $k=1$. It is also true for $k=2$ by virtue of the commutative and idempotent properties of $\mathscr{P}$ in equation (16). The proof follows directly by the method of induction. □

Lemma 3: $[\mathbf{P}-\mathscr{P}]^k = \mathbf{P}^k - \mathscr{P}\ \ \forall k \in \mathbb{N}$.

Proof: The proof follows by setting $\alpha = 1$ in Proposition 4. □

Proposition 5:

$$\lim_{\theta\to 0^+}\theta[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1} = \mathscr{P}. \tag{18}$$

Proof: For $\theta \in (0,1)$, it follows from equation (16) and Lemma 3 that

$$\begin{aligned}&\theta[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1}-\mathscr{P}\\ &\quad= \theta\sum_{k=0}^{\infty}\left((1-\theta)^k\mathbf{P}^k\right) - \theta\sum_{k=0}^{\infty}(1-\theta)^k\ \mathscr{P}\\ &\quad= \theta\sum_{k=0}^{\infty}(1-\theta)^k\left(\mathbf{P}^k-\mathscr{P}\right)\\ &\quad= \theta\sum_{k=0}^{\infty}(1-\theta)^k\left(\mathbf{P}-\mathscr{P}\right)^k \text{ by Lemma 3}\\ &\quad= \theta[\mathbb{I}-(1-\theta)(\mathbf{P}-\mathscr{P})]^{-1} \text{ by Lemma 2}\\ &\quad\Rightarrow \lim_{\theta\to 0^+}\left(\theta[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1}-\mathscr{P}\right)\\ &\quad= \lim_{\theta\to 0^+}\theta[\mathbb{I}-(1-\theta)(\mathbf{P}-\mathscr{P})]^{-1}.\end{aligned}$$

Since, for continuous functions $f(\cdot)$ and $g(\cdot)$ with

$$\begin{aligned}&\lim_{\theta\to 0^+}f(\theta) = 0 \quad \text{and} \quad \lim_{\theta\to 0^+}g(\theta) < \infty\\ &\Longrightarrow \lim_{\theta\to 0^+}f(\theta)g(\theta) = 0,\end{aligned}$$

it follows from Lemma 2 that

$$\begin{aligned}&\lim_{\theta\to 0^+}\theta[\mathbb{I}-(1-\theta)(\mathbf{P}-\mathscr{P})]^{-1} = 0\\ &\quad\Rightarrow \lim_{\theta\to 0^+}\left(\theta[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1}-\mathscr{P}\right) = 0.\end{aligned}$$

The proof is thus complete. □

Proposition 6: *For every stochastic matrix* $\mathbf{P}$, *the generalized measure is expressed as*

$$\boldsymbol{\vartheta}^0|_p \equiv \lim_{\theta\to 0^+}\boldsymbol{\vartheta}^\theta|_p = \left\{\begin{array}{c}\vdots\\ \dfrac{\mathscr{P}_{i\circ}\boldsymbol{\chi}}{\|\mathscr{P}_{i\circ}\|_p}\\ \vdots\end{array}\right\}, \tag{19}$$

where $\mathscr{P}_{i\circ}$ *is the ith row of* $\mathscr{P}$.

Proof: Following equation (11) in Definition 9, it suffices to show that

$$\lim_{\theta\to 0}\begin{bmatrix} \xi_1^p(\theta) & \cdots & 0 \\ \vdots & \xi_i^p(\theta) & \vdots \\ 0 & \cdots & \xi_n^p(\theta) \end{bmatrix}[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1} = \begin{bmatrix} \|\mathscr{P}_{1\circ}\|_p & \cdots & 0 \\ \vdots & \|\mathscr{P}_{i\circ}\|_p & \vdots \\ 0 & \cdots & \|\mathscr{P}_{n\circ}\|_p \end{bmatrix}^{-1}\begin{bmatrix} \mathscr{P}_{1\circ} \\ \cdots \\ \mathscr{P}_{n\circ} \end{bmatrix}$$

The above identity is a direct consequence of the following two relations:

$$\lim_{\theta\to 0}\theta\big[\mathbb{I}-(1-\theta)\mathbf{P}\big]^{-1} = \mathscr{P}$$

$$\lim_{\theta\to 0^+}\theta^{-1}\begin{bmatrix} \xi_1^p(\theta) & \cdots & 0 \\ \vdots & \xi_i^p(\theta) & \vdots \\ 0 & \cdots & \xi_n^p(\theta) \end{bmatrix} = \begin{bmatrix} \|\mathscr{P}_{1\circ}\|_p & \cdots & 0 \\ \vdots & \|\mathscr{P}_{i\circ}\|_p & \vdots \\ 0 & \cdots & \|\mathscr{P}_{n\circ}\|_p \end{bmatrix}^{-1}.$$

The first relation is a restatement of equation (18) in Proposition 6. The second relation is obtained from continuity of norm in equation (11) (see Remark 1). □

We now consider the special class of primitive (i.e., irreducible and acyclic (Bapat and Raghavan 1997)) stochastic matrices. The restriction of primitivity is valid for many applications such as finite-state machines without any deadlock or local livelock. A primitive stochastic matrix **P** has the following properties (Bapat and Raghavan 1997):

(i) $\lim_{k\to\infty}\mathbf{P}^k = \mathscr{P}$ and $\mathbf{P}\mathscr{P} = \mathscr{P}\mathbf{P} = \mathscr{P} = \mathscr{P}^2$

(ii) The matrix $\mathscr{P}$ has the following structure:

$$\mathscr{P} = \begin{bmatrix} \wp^T \\ \cdots \\ \wp^T \end{bmatrix} \quad \text{where } \wp^T\mathbf{P} = \wp^T$$

implying that $\wp$ is the left eigenvector of **P** corresponding to its unique unity eigenvalue

(iii) Upon ℓ_1-normalization, $\wp$ becomes the state probability vector of the stationary Markov chain associated with the stochastic primitive matrix **P**.

(iv) The spectral radius of the matrix $(\mathbf{P}-\mathscr{P})$ is less than unity, i.e., the eigenvalues of $(\mathbf{P}-\mathscr{P})$ are located within the unit radius circle with center at the origin.

For a primitive stochastic matrix, the expression for $\boldsymbol{\vartheta}^0|_p$ in equation (19) of Proposition 6 is simplified as presented in the following proposition.

Proposition 7: *For a primitive stochastic matrix* **P**, *the generalized measure is expressed as*

$$\boldsymbol{\vartheta}^0|_p \equiv \lim_{\theta\to 0^+}\boldsymbol{\vartheta}^\theta|_p = \frac{\wp^T\boldsymbol{\chi}}{\|\wp\|_p}\begin{Bmatrix} 1 \\ \vdots \\ 1 \end{Bmatrix}, \tag{20}$$

where $\wp^T\mathbf{P} = \wp^T$.

Proof: From the properties (i) and (ii) of primitive matrices, it follows that

$$\mathscr{P}_{j\circ} = \wp^T \ \forall j \in \{1,\ldots,n\}, \tag{21}$$

where $\wp^T$ is the state probability vector of the associated Markov chain. Then, the proof follows from Proposition 6. □

3.2 *Physical interpretation of the* $\boldsymbol{\vartheta}^0|_p$ *measures*

All entries of the $\boldsymbol{\vartheta}^0|_p$ vector in equation (19) are identical for a primitive stochastic matrix and hence a single entry can be taken as a scalar measure, $\vartheta^0|_p \equiv \wp^T\chi/\|\wp\|_p$, of the regular language of the underlying automaton. For all $p \in [1,\infty]$, the measure $\vartheta^0|_p$ represents the long-range behaviour of the plant dynamics in terms of the (assigned) terminating characteristics and the stationary state probability vector of the finite Markov chain model. However, the measures for different values of p are not equivalent in the sense that a control policy optimizing $\vartheta^0|_p$ does not necessarily coincide with one that optimizes $\vartheta^0|_q$ for $p \neq q$. For example, a control policy that maximizes $\vartheta^0|_1$ selectively disables controllable events such that $\wp^T\boldsymbol{\chi}$ is maximized; and a control policy that maximizes $\vartheta^0|_2$ chooses an automaton configuration to make the stationary state probability vector $\wp$ closest to the terminal characteristic vector $\boldsymbol{\chi}$ in the Euclidean sense. For physical understanding and visualization, let $\mathscr{S}$ be a bounded submanifold of $\mathbb{R}^n$ such that

$$\forall\, p = \{p_1,\ldots,p_n\} \in \mathscr{S} \quad \text{with} \begin{cases} p_i \geq 0 \\ \sum_{i=0}^n p_i = \|p\|_{l_1} = 1. \end{cases} \tag{22}$$

Then, for any n-state automaton, the stationary state probability vector is $\wp \in \mathscr{S}$. Figure 2 illustrates $\mathscr{S}$ for

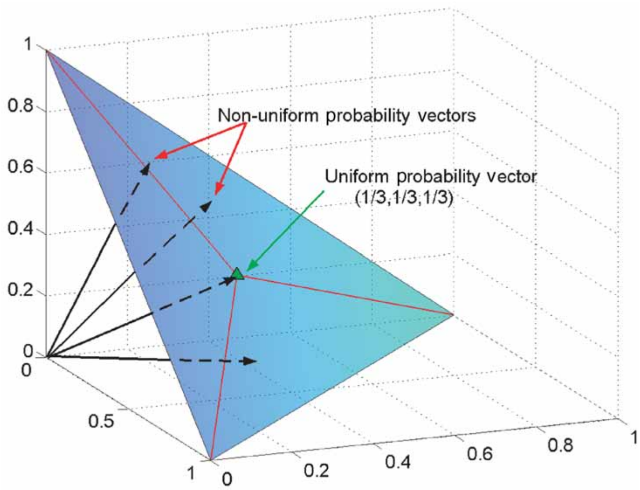

Figure 2. Representation of the $\mathscr{S}$-plane for three states.

Lemma 4: *Let* ν *be a* n*-dimensional vector with* $\nu_i \in [0,1]$ *and* $\sum_i \nu_i = 1$. *Then we have*

$$\text{(i)} \quad \|\nu\|_p \in \left[n^{(1-p)/p}, 1\right] \quad \forall p \in [1, \infty] \tag{23}$$

$$\text{(ii)} \quad \|\nu\|_p \le \|\nu\|_q \quad \forall p > q \text{ with } p, q \in [1, \infty] \tag{24}$$

Proof: For Assertion (i),

$$\nu_i \in [0,1] \;\Rightarrow\; \nu_i^p \le \nu_i \;\Rightarrow\; \sum_i \nu_i^p \le \sum_i \nu_i \;\Rightarrow\; \|\nu\|_p \le 1.$$

The result follows by noting that the smallest value is attained when all ν_i are equal i.e. $\nu_i = 1/n \; \forall i$. Assertion (ii) follows by noting that $\nu_i^p \le \nu_i^q$ if $p > q$. □

$n = 3$, where the central point $\wp_c$ denotes the uniform probability vector $[1/n, \ldots, 1/n]$, which is interpreted to have maximum entropy $\log_2 n$ in the Shannon sense (Cover and Thomas 1991). Moving away from $\wp_c$ on the $\mathscr{S}$-plane, the distribution becomes non-uniform, i.e., the Shannon entropy $S \equiv -\sum_{k=1}^{n} \left(p_k \log_2 p_k\right)$ decreases toward zero.

In view of the above discussion and Lemma 4, a supervisory control policy can be constructed by optimizing $\vartheta^0|_p$ in the sense of equation (19) for a specified $p \in [1, \infty]$ to obtain a stationary state probability vector $\wp$. For example, if $p = 1$, then the optimization algorithm attempts to choose $\wp$ as a unit vector in the direction of one of the axes of $\mathbb{R}^n$ for which the $\boldsymbol{\chi}$-vector has the largest element; if this is the case, then Shannon entropy $S = 0$. If $p = 2$, then the optimization algorithm attempts to choose $\wp$ as the point of intersection of the $\boldsymbol{\chi}$ vector with the $\mathscr{S}$-plane; in this case, the Shannon entropy is $S > 0$ unless $\boldsymbol{\chi}$ is coincident with one of the axes of $\mathbb{R}^n$. For $p > 2$, the algorithm attempts to choose $\wp$ closer to the central point $\wp_c$ more and more aggressively as p increases toward infinity, for which the Shannon entropy S increases toward its maximum value $\log_2 n$.

Three measures are considered to be significant; $\vartheta^0|_1$ and $\vartheta^0|_\infty$ optimal policies are useful to obtain low and high entropy (thermodynamically stable) distributions, respectively, and $\vartheta^0|_2$ optimality is useful when the problem definition requires achieving a target distribution over the plant states as closely as the controllability criteria would allow. An example is given in § 5.

The following lemma is useful for interpretation of the $\boldsymbol{\vartheta}^0|_p$ measures.

4. Shaped measures

The measures defined in the previous sections put equal importance to all traces in the generated language of an automaton, including traces of unbounded lengths. This section investigates formal measures that generalize the process of assigning importance or weight to a trace as a function of its length. For a given automaton G_i, a partition of the generated language $L(G_i)$ is obtained as:

$$L(G_i) = \bigcup_{i=0}^{\infty} \mathscr{L}_i^r \quad \text{where} \quad \mathscr{L}_i^r = \left\{\omega \in L(q_i, q_j) : \; |\omega| = r\right\}. \tag{25}$$

(Note that $\mathscr{L}_i^r \bigcap_{r \neq s} \mathscr{L}_i^s = \emptyset$.) From σ-additivity of the language measure (Ray 2005), the following notion of language measure is introduced.

Definition 10: For a given starting state q_i and the parameter $\theta \in (0,1)$, the shaped measure of the language $L(G_i)$ is defined as

$$\nu(L(G_i)) = \sum_{r=0}^{\infty} \mu_i^{\theta}(\mathscr{L}_i^r). \tag{26}$$

The above definition, as observed before, fails to exist as $\theta \to 0^+$. The singularity at $\theta = 0$ is alleviated by a shaping sequence that provides appropriate weights on the individual terms of the infinite sum in equation (26). The next proposition establishes that every ℓ_1-sequence qualifies as a shaping sequence.

Proposition 8: *For a real ℓ_1-sequence $\Gamma = \{\gamma_i\}$ with $\gamma_i \in [0, \infty)$,*

$$\text{(i)}\ \sum_{i=0}^{\infty} \mu_i^{\theta}(\mathscr{L}^i)\gamma_i < \infty \quad \forall\, \theta \in (0, 1) \tag{27}$$

$$\text{(ii)}\ \lim_{\theta \to 0^+} \sum_{i=0}^{\infty} \mu_i^{\theta}(\mathscr{L}^i)\ \gamma_i < \infty. \tag{28}$$

Proof: The proof of the proposition requires the following lemma. □

Lemma 5: *The following expression holds for the θ-parameterized shaped measure*

$$\boldsymbol{\mu}^{\theta}(\mathscr{L}_i^r) = (1-\theta)^r \mathbf{P}^r \boldsymbol{\chi}. \tag{29}$$

Proof: From Definition 7, we have

$$\begin{aligned} \mu_i^{\theta}(\mathscr{L}_i^r) &= \sum_{j=1}^{n} \sum_{\substack{\omega \in \\ L(q_i, q_j) \cap \mathscr{L}_i^r}} \tilde{\pi}(q_i, \omega)\chi_j \\ &= \sum_{j=1}^{n} \left\{ \sum_{i_1} \cdots \sum_{i_r} \pi_{ii_1} \cdots \pi_{i_r j} \right\} \chi_j \\ &= \sum_j \Pi_{ij}^r \chi_j = \sum_j (1-\theta)^r \mathbf{P}_{ij}^r \chi_j \end{aligned}$$

Using Lemma 5 and noting that for all $\theta \in [0, 1)$, it follows that

$$\|(1-\theta)^r \mathbf{P}^r \boldsymbol{\chi}\|_{\infty} \le |1-\theta|^r\ \|\mathbf{P}\|_{\infty}^r\ \|\boldsymbol{\chi}\|_{\infty} \le 1, \tag{30}$$

where $\| \bullet \|_{\infty}$ is the induced sup-norm of $\bullet$. □

The shaped measure is now formally defined based on Proposition 33 by setting the parameter θ to 0.

Definition 11: Let $\Gamma = \{\gamma_i\}$ be a ℓ_1-sequence of non-negative real numbers (called the shaping sequence in the sequel). The shaped measure τ_i^{Γ} of a trace set $\{s\} \subseteq L(q_i, q_j) \subseteq L(G_i)$ with $s = k \in \mathbb{N} \cup \{0\}$ relative to Γ is defined as

$$\tau_i^{\Gamma}(\{s\}) \equiv \mu^{\theta}(\{s\})_k = \tilde{\pi}(s, q_i)\chi(q_j)_k \quad \forall s \in L(q_i, q_j). \tag{31}$$

The shaped measure of $L(q_i, q_j)$ is defined as:

$$\tau_i^{\Gamma}\big(L(q_i, q_j)\big) \equiv \sum_{r=0}^{\infty} \sum_{\substack{\omega \in \\ L(q_i, q_j) \cap \mathscr{L}_i^r}} \tau_i^{\Gamma}(\{s\}). \tag{32}$$

The shaped measure of a *DFSA* G_i, relative to the sequence Γ and initialized at the state $q_i \in Q$, is denoted as: $\tau_i^{\Gamma} \equiv \tau_i^{\Gamma}(L(G_i)) = \sum_j \tau_i^{\Gamma}(L(q_i, q_j))$.

The shaped measure vector, relative to the sequence Γ, is denoted as: $\tau^{\Gamma} \equiv [\tau_1^{\Gamma}, \ldots, \tau_n^{\Gamma}]$.

Remark 2: If the short-term behaviour of the discrete-event system is of interest, then all but finitely many elements of the shaping sequence $\Gamma = \{\gamma_i\}$ in Definition 11 could be restricted to be zeros. Then, there exists $r^\star \in \mathbb{N}$ such that $\mathscr{L}_i^r = \emptyset\ \forall r \ge r^\star$, i.e., the generated language has only bounded length traces.

4.1 *Relation between $\tau^{\Gamma(p)}$ and $\boldsymbol{\vartheta}^0|_p$ measures*

In spite of a different construction, shaped measures are related to the generalized measure defined in §3. Specifically, there exist sequences of shaped measures that converge to $\boldsymbol{\vartheta}^0|_p$.

Remark 3: Let $p \in [1, \infty]$ and let $\Gamma_k(p), k \in \mathbb{N}$ be a sequence of non-negative real numbers, whose all elements, except the kth one, are zeroes and the kth element is $\|\wp\|_p$. Let $\Gamma(p) \equiv \lim_{k\to\infty} \Gamma_k(p)$. Then, it follows from Proposition 6 or Proposition 7 that there exists $\Gamma(p)$ such that $\tau^{\Gamma(p)} = \boldsymbol{\vartheta}^0|_p\ \forall p \in [1, \infty]$.

4.2 *Physical interpretation of shaped measures*

A shaping sequence Γ specifies length-based relative importance of traces in the generated language. Intuitively, one is rarely interested in all traces generated by an automaton. More often than not, either short traces or very long traces (specifically of unbounded length) are important. The first case is handled by shaping sequences with finitely many non-zero terms and the latter, shown in Remark 3 is viewed as a limit of the shaped measures. However, shaping sequences can be more complicated; the only requirement is that the sequence be in ℓ_1 (see Proposition 8). In this context, Remark 3 implies that $\vartheta^0|_1$ addresses the long-term behaviour of the discrete-event system based on the traces of unbounded length with no importance to finite traces. This follows from the fact that, for $p = 1$ and all elements of the sequence Γ_k are zeros with the exception of the k^{th} element being equal to 1.

5. An illustrative example

Figure 3 shows the finite-state automaton model of the plant, where the state set $Q = \{q_1, \ldots, q_9\}$ and the event alphabet $\Sigma = \{\sigma_r, \sigma_l, \sigma_f, \sigma_b, \sigma_{fl}, \sigma_{rf}, \sigma_{rb}, \sigma_{lb}, \omega_1\}$. The transitions, shown by dashed lines, are controllable and those, shown by solid lines, are uncontrollable. The state transition matrix $\mathbf{P}$ is given in table 1. The stochastic matrix $\mathbf{P}$ is primitive because $\mathbf{P}^2$ is a positive matrix. The stationary state probability vector of $\mathbf{P}$ and the

χ-vector for the *DFSA* are

$$\wp^T = \begin{bmatrix} 0.234 & 0.041 & 0.065 & 0.084 & 0.095 & 0.016 & 0.153 & 0.147 & 0.076 \end{bmatrix}$$
$$\chi = \begin{bmatrix} 0.66 & -0.42 & -0.97 & 0.52 & -0.49 & -0.57 & 0.57 & -0.09 & 0.43 \end{bmatrix}^T.$$

The scalar measures $\vartheta^0|_p$ and $\|\wp\|_p$ for the primitive matrix in table 1 are plotted for different values of p in figure 4. A shaping sequence Γ evaluates the short-term behaviour based on traces of length less than 60 and the resulting measure vector is

$$\tau^\Gamma = \begin{bmatrix} 0.276 & -0.252 & 0.627 & -0.375 & 0.256 & -0.055 & -0.059 & -0.103 & 0.475 \end{bmatrix}^T.$$

Supervisory control policies have been computed based on $\vartheta^0|_p$, $\|\wp\|_p$ and τ^Γ by optimizing the respective scalar measures via a standard search algorithm. Different values of $p = 1, 4$, and ∞ are chosen to illustrate the fact that they result in different optimal control policies. The choices of $p=1$ and $p=\infty$ are made as they are familiar norms used in engineering analysis; and the choice of $p=4$ is made because it effects are intermediate between $p=1$ and $p=\infty$ and are different from those of the Euclidean norm $p=2$ (see §3). For these cases, the results of improved performance under optimal supervision are summarized below.

For $p=1$, $\vartheta^0|_1$ is increased from 0.12 to 0.35.
For $p=4$, $\vartheta^0|_4$ is increased from 0.48 to 1.03.
For $p=\infty$, $\vartheta^0|_\infty$ is increased from 0.52 to 1.3 and τ^Γ is increased elementwise from

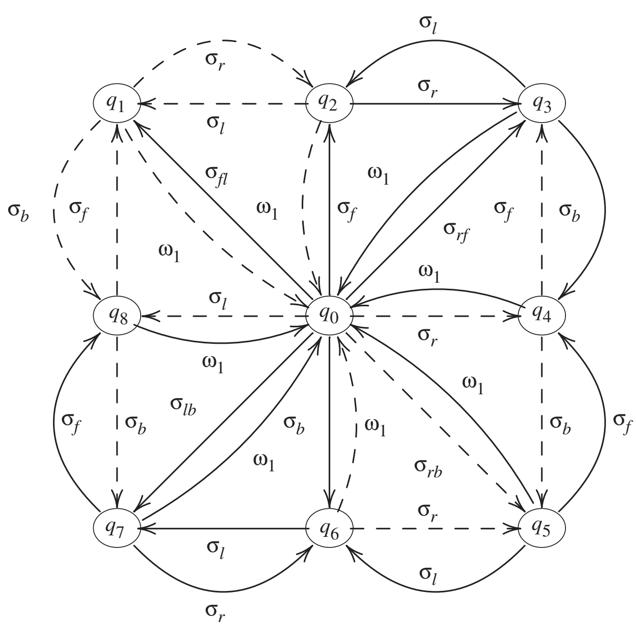

Figure 3. Plant automaton model.

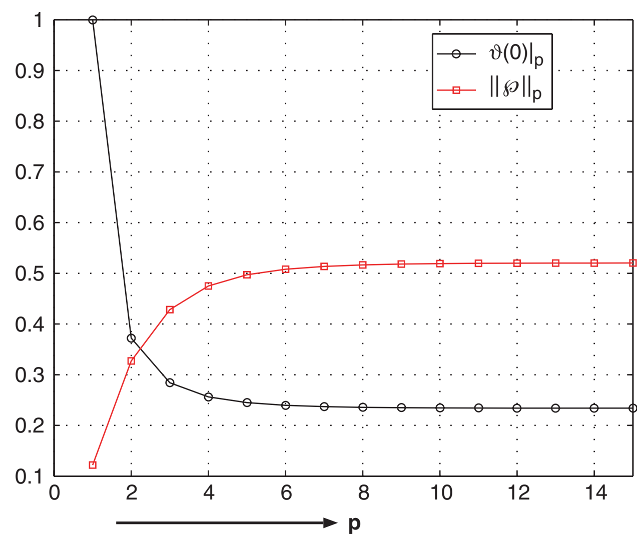

Figure 4. Profiles of $\vartheta^0|_p$ and $\|\wp\|_p$ with p. Note the stabilizing feature of each plot for increasing p.

Table 1. State transition matrix **P** of the plant automaton model.

0	0.015	0.102	0.041	0.120	0.048	0.300	0.139	0.139
0.372	0	0.131	0	0	0	0	0	0
0.130	0.319	0	0.551	0	0	0	0	0
0.087	0	0.424	0	0.489	0	0	0	0
0.351	0	0	0.411	0	0.238	0	0	0
0.337	0	0	0	0.240	0	0.423	0	0
0.069	0	0	0	0	0.470	0	0.460	0.460
0.738	0	0	0	0	0	0.259	0	0
0.199	0.218	0	0	0	0	0	0.583	0.583

Table 2 Optimal decision for disabling of controllable events.

Controllable events	$\vartheta^0\|_1$	$\vartheta^0\|_4$	$\vartheta^0\|_\infty$	τ^Γ
$q_1 \overset{\sigma_r}{\rightarrow} q_5$	×	×	×	×
$q_1 \overset{\sigma_{rb}}{\rightarrow} q_6$	×	×	×	×
$q_1 \overset{\sigma_l}{\rightarrow} q_9$	✓	✓	✓	✓
$q_2 \overset{\omega_1}{\rightarrow} q_1$	✓	✓	✓	✓
$q_2 \overset{\sigma_r}{\rightarrow} q_3$	×	×	×	×
$q_2 \overset{\sigma_b}{\rightarrow} q_9$	✓	×	✓	×
$q_3 \overset{\omega_1}{\rightarrow} q_1$	✓	✓	✓	✓
$q_5 \overset{\sigma_f}{\rightarrow} q_4$	✓	✓	✓	✓
$q_5 \overset{\sigma_b}{\rightarrow} q_6$	✓	✓	✓	✓
$q_7 \overset{\omega_1}{\rightarrow} q_1$	×	×	×	✓
$q_7 \overset{\sigma_r}{\rightarrow} q_6$	×	×	×	×
$q_9 \overset{\sigma_b}{\rightarrow} q_8$	×	×	×	×
$q_9 \overset{\sigma_f}{\rightarrow} q_2$	×	×	✓	×

$[1.2 \;\; 1.5 \;\; 0.9 \;\; 0.3 \;\; 1.1 \;\; 1.4 \;\; 1.2 \;\; 1.7 \;\; 1.4]^T$ to $[3.4 \;\; 3.8 \;\; 2.5 \;\; 1.7 \;\; 2.6 \;\; 3.3 \;\; 3.5 \;\; 3.7 \;\; 4.0]^T$.

Table 2 enumerates the optimal decision sets for disabling of controllable events obtained in the above four cases, where × and ✓ indicate disabled controllable events and enabled controllable events, respectively. The decisions are made from the stationary state probability distributions achieved from the optimal policies shown in figure 5. It is seen that the $\vartheta^0|_1$-optimal policy achieves both maximum and minimum probability values in states 9 and 6, respectively. This shows that $\vartheta^0|_1$-optimal policy does indeed produce relatively less uniform distribution in comparison to the $\vartheta^0|_\infty$-optimal policy, as stated in § 3. It is also noted that the τ^Γ-optimal policy achieves the least uniform distribution.

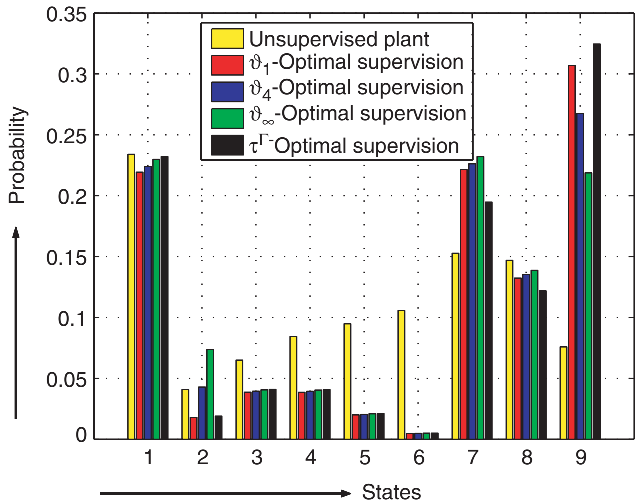

Figure 5. Stable distributions for computed optimal policies.

6. Summary, conclusions, and future work

This paper formulates and validates a concept of generalization of signed real measure of regular languages, which also leads to renormalization (Chattopadhyay and Ray 2006) of the normalized measure and eliminates the need for a user-selectable parameter in the original concept of language measure Ray (2005). These generalizations are achieved through a trace measure that is characterized by both initial and terminal states as well as the length of the trace and the choice of a vector norm for renormalization. The generalized measures with different norms are not equivalent in the sense that the respective optimal control policies with these measures as the performance cost functionals are different. These concepts are illustrated with simple examples for quantitative analysis and synthesis of discrete-event supervisory control systems. It is envisioned that optimal supervisory decision & control of discrete-event systems (Sengupta and Lafortune 1998, Ray *et al*. 2004) can be enhanced through appropriate selection of a language measure to enhance the objectives at hand. In this context, future research is recommended in the following areas:

- Generalization of the language-measure-based optimal control algorithms (Ray *et al*. 2004) for sub-stochastic transition matrices to the stochastic case. A potential application is to compute a sufficiently small termination probability θ such that, as $\theta \rightarrow 0^+$, the optimal control policies approach the true situation for the non-terminating plant.
- Extension the concept of (regular) language measure for (non-regular) languages higher up in the Chomsky Hierarchy such as context-free and context -sensitive languages. A first attempt to extend the concept of the language measure to linear grammars was reported in (Ray *et al*. 2004). Further investigations in this direction is required for extension of the concept to more complex models.
- Applications of language measure in anomaly detection, model identification and order reduction, and construction of interfaces between continuously varying and discrete-event spaces.

Acknowledgements

This work has been supported in part by the U.S. Army Research laboratory and the U.S. Army Research Office under Grant Nos. DAAD19-01-1-0646 and W911NF-06-1-0469.

References

R. Bapat and T. Raghavan, *Nonnegative Matrices and Applications*, Cambridge, UK: Cambridge University Press, 1997.

I. Chattopadhyay and A. Ray, "A complex measure for linear grammars", *Demonstratio Mathematica*, 38, pp. 761–775, 2005.

I. Chattopadhyay and A. Ray, "Renormalized measure of regular languages", in *Proceedings of WODES 2006*, submitted for review, (Ann Arbor, MI), July 2006.

T. Cover and J. Thomas, *Elements of Information Theory*, New York: John Wiley, 1991, pp. 18–19.

V. Garg, "An algebraic approach to modeling probabilistic discrete event systems", in *Proceedings of 1992 IEEE Conference on Decision and Control*, Tucson, AZ, December 1992a, pp. 2348–2353.

V. Garg, "Probabilistic lnaguages for modeling of DEDs", in *Proceedings of 1992 IEEE Conference on In formation and Sciences*, Princeton, NJ, March 1992, pp. 198–203.

J. Hopcroft, R. Motwani and J. Ullman, *Introduction to Automata Theory, Languages, and Computation*, 2nd ed., Boston, MA, USA: Addison-Wesley, 2001.

A. Naylor and G. Sell, *Linear Operator Theory in Engineering and Science*, New York, USA: Springer-Verlag, 1982.

P. Ramadge and W. Wonham, "Supervisory control of a class of discrete event processes", *SIAM J. Cont. Optim.*, 25, pp. 206–230, 1987.

A. Ray, "Signed real measure of regular languages for discrete-event supervisory control", *Int. J. Cont.*, 78, pp. 949–967, 2005.

A. Ray, J. Fu and C. Lagoa, "Optimal supervisory control of finite state automata", *Int. J. Cont.*, 77, pp. 1083–1100, 2004.

A. Ray, V. Phoha and S. Phoha, *Quantitative Measure for Discrete Event Supervisory Control*, New York: Springer, 2005.

R. Sengupta and S. Lafortune, "An optimal control theory for discrete event systems", *SIAM J. Cont. Optim.*, 36, pp. 488–541, 1998.