Renormalized measure of regular languages

I. CHATTOPADHYAY and A. RAY*

The Pennsylvania State University, University Park, PA 16802

This paper modifies the signed real measure of regular languages, which has been reported in recent literature for analysis and synthesis of discrete event supervisory control laws. A new concept of renormalized measure is introduced to eliminate a user-selectable parameter in the present version of the language measure. The concept of measure renormalization is illustrated by an example.

1. Introduction

The signed real measure of regular languages (of symbolic strings) has been reported in recent literature (Ray 2005, Ray *et al.* 2005) as a quantative tool for analysis and synthesis of discrete event supervisory control laws (Ray *et al.* 2004). The regular language of plant dynamics is realized by a Deterministic Finite State Automaton (DFSA) $G_i \equiv \langle Q, \Sigma, \delta, q_i, Q_m \rangle$ with cardinality of the set of states $|Q| = n \in \mathbb{N}$. The language measure is obtained in terms of a normalized state weight vector, called the χ-vector, and the state transition cost matrix, called the a Π-matrix. While the elements of the χ-vector are normalized to lie in the interval $[-1, 1]$ and the reflect individual state's relative impact on the system performance, the Π-matrix captures the dynamics behaviours of the plant under control. An element π_{jk} of the Π-matrix is conceptually similar to the state transition probability of Markov Chain with the exception that the equality condition $\Sigma_k \pi_{jk} = 1$ is not satisfied. Specifically, $\Sigma_k \pi_{jk} < 1, j = 1, 2, \ldots, n$ makes Π a contraction operator; this strict inequality is based on the following rationale (Ray 2005, Ray *et al.* 2005). Since the plant model is an inexact representation of the physical plant, there exist unmodelled dynamics to be accounted for. This phenomenon manifests itself either as unmodelled events that may occur at each state or as unaccounted state(s) in the model.

Garg *et al.* (1992a, b) studied probabilistic discrete event systems in the context of "terminating" models, where the event probabilities at one or more state may not sum to unity. The difference from unity is interpreted as the probability of the plant terminating operation at the particular state. Equivalently, the difference from unity is the probability of transition to a dump state. Reported work on language measure (Ray 2005) and corresponding optimal control algorithms (Ray *et al.* 2004) investigated such "terminating" plant models.

Let Σ^u denote the set of all unmodelled events in the DFSA. A new unmarked absorbing state q_{n+1}, called the dump state (Ramadge and Wonham 1987), is created and the transition function δ is extented to δ_{ext}: $(Q \cup \{q_{n+1}\}) \times (\Sigma \cup \Sigma^u) \rightarrow (Q \cup \{q_{n+1}\})$. The residue $\theta_j = 1 - \Sigma_k \pi_{jk}$ denotes the probability of transition from q_j to q_{n+1}. Consequently, the Π-matrix is augmented to obtain a stochastic matrix Π_{aug} as follows:

$$\Pi_{\text{aug}} = \begin{bmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1n} & \theta_1 \\ \pi_{21} & \pi_{22} & \cdots & \pi_{2n} & \theta_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \pi_{n1} & \pi_{n2} & \cdots & \pi_{nn} & \theta_n \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}. \quad (1)$$

Since the dump state q_{n+1} is not marked (Ramadge and Wonham 1987), it follows from (Ray 2005) that the corresponding state weight $\chi_{n+1} = 0$. Hence, the χ-vector is augmented as $\chi_{\text{aug}} = [\chi^T \ 0]^T$. With these extension, the language measure vector $\mu_{\text{aug}} \equiv [\mu_1 \ \mu_2 \ \ldots \ \mu_n \ \mu_{n+1}]^T = [\mu^T \ \mu_{n+1}]^T$ of the

*Corresponding author. Email: axr2@psu.edu

augmented DFSA is expressed as

$$\mu_{\text{aug}} \equiv \begin{pmatrix} \mu \\ \mu_{n+1} \end{pmatrix} = \begin{pmatrix} \Pi\mu + [\theta_1 \cdots \theta_n]^T \mu_{n+1} \\ \mu_{n+1} \end{pmatrix} + \begin{pmatrix} \chi \\ 0 \end{pmatrix}. \tag{2}$$

In equation (2), $\mu_{n+1} = 0$ because $\chi_{n+1} = 0$ and all transitions from the absorbing state q_{n+1} lead to itself. Hence,

$$\mu(\theta_1, \ldots, \theta_n) = \Pi(\theta_1, \ldots, \theta_n)\mu(\theta_1, \ldots, \theta_n) + \chi. \tag{3}$$

Since it may be difficult to identify θ_j, $j = 1, 2, \ldots, n$ from experimental data, these n parameters are replaced by a single parameter $\theta \equiv \min_j \theta_j$ such that the induced sup-norm $\|\Pi\|_\infty = \max_j(1 - \theta_j)(\theta)$ remains unchanged. In that case, $\Pi(\theta)$ can be obtained by multiplying the experimentally determined stochastic matrix of state transitions with the scalar parameter $(1 - \theta)$ (Ray 2005, Ray *et al.* 2005). The objective here is to formulate a well-defined measure for the limiting non-terminating case. Therefore, identical probabilities of transition from each terminating state to the dump state (i.e., a common θ) does not cause any loss of generality. The next section shows that this approach makes the computed measures unique.

This paper introduces the new concept of renormalized measure and its properties to eliminate the parameter θ in the language measure. Specifically, the renormalized measure is computed in the limit θ approaching zero so that $\Pi(\theta)$ converges to the original (i.e., experimentally determined) stochastic matrix (Ray 2005, Ray *et al.* 2005) of state transitions. Specifically, the language measure is extended to "non- terminating" plants (i.e., to automata that do not have a dump state).

This paper is organized in five sections and appendix. Section 2 introduces the concept of renormalized measure and derives its properties. Section 3 computes recursive relations on the derivatives of the renormalized measure and establishes their limits. Section 4 presents an example to illustrate the concept of renormalized measure. Section 5 summarizes and concludes the paper with recommendations for future work. The appendix presents pertinent results on stochastic matrices to support §§ 2 and 3.

2. Renormalization of language measure

This section introduces the concept of renormalized measure as an extension of the signed real measure μ (Ray 2005, Ray *et al.* 2005). Let $\mathbf{P} \in [0, 1]^{n \times n}$ be a stochastic matrix (Bepat and Raghavan 1997) (i.e., each element of $\mathbf{P}$ is non negative and each row of $\mathbf{P}$ sums to unity), implying that the induced sup-norm $\|\mathbf{P}\|_\infty = 1$. Let us now parameterize the matrix operator $\mathbf{P}$ in terms

of $\theta \in (0, 1)$ as

$$\Pi(\theta) \equiv (1 - \theta)\mathbf{P}. \tag{4}$$

Then, the induced sup-norm $\|\Pi(\theta)\|_\infty = (1 - \theta)$. Let us consider a regular language and its realization as a deterministic finite state automaton (DFSA) over a symbol alphabet Σ, where the number of automaton states is $n \in \mathbb{N}$. Following equations (2)–(4), the language measure is obtained (Ray 2005, Ray *et al.* 2005) as

$$\mu(\theta) = [I - \Pi(\theta)]^{-1}\chi, \tag{5}$$

where $\chi \in [-1, 1]^n$ is the normalized state weight vector for the DFSA; and $\Pi(\theta)$ is the state transition cost matrix (Ray 2005). Since $\Pi(\theta)$ is a linear continuous operator in the Banach space $\mathbb{R}^n$ with the induced sup-norm $\|\Pi(\theta)\|_\infty = 1 - \theta$, the following inequality (Naylor and sell 1982) holds:

$$\|[I - \Pi(\theta)]^{-1}\|_\infty \leq \theta^{-1}. \tag{6}$$

Furthermore, since $\|\chi\|_\infty \leq 1$, it follows from equations (5) and (6) that

$$\|\mu(\theta)\|_\infty \leq \theta^{-1} \quad \forall \, \theta \in (0, 1). \tag{7}$$

The regular language under consideration is a sub-language of the Kleene closure Σ^* of the alphabet Σ, for which the automation states can be merged into a single state. In that case, the state transition cost matrix $\Pi(\theta)$ degenerates to the 1×1 matrix $[1 - \theta]$ and the normalized state weight vector χ becomes one-dimensional and can be assigned as $\chi = 1$; consequently, the measure vector $\mu(\theta)$ degenerates to the scalar measure θ^{-1}. Hence, to alleviate the singularity of the matrix operator $[I - \Pi(\theta)]$ at $\theta = 0$, the measure vector $\mu(\theta)$ in equation (5) is normalized with respect to θ^{-1} to obtain the renormalized measure vector.

Definition 1:

$$\mu^{\text{ren}}(\theta) \equiv \theta[I - \Pi(\theta)]^{-1}\chi$$
$$= \theta[I - \mathbf{P}(1 - \theta)]^{-1}\chi \tag{8}$$

Remark 1: Recalling that the original expression for language measure in equation (5) is given by $[I - \Pi]^{-1}\chi$, the transition probability matrix Π becomes a stochastic matrix (i.e., each row summing exactly to unity) as $\theta \to 0^+$; hence $[I - \Pi]$ becomes non-invertible. It is shown in the following proposition that the problem can be handled by considering non terminating models as the limiting cases of terminating models as the transition probabilities to the dump state (see § 1) goes to zero.

Propositon 1: *The limiting renormalized measure vector $\mu^{\text{ren}}(0) \equiv \lim_{\theta \to 0+} \mu^{\text{ren}}(\theta)$ exists and $\|\mu^{\text{ren}}(0)\|_\infty \leq 1$.*

Proof: Because of continuity of the linear bounded operator Π and completeness of the space $\mathbb{R}^n$, equations (6)–(8) yield the following result: $\|\mu^{\text{ren}}(0)\|_\infty \leq 1$. $\qquad\square$

Proposition 2: *Let P be the transition matrix of a finite markov chain (or equivalently a regular language). Then, as the parameter $\theta \to 0^+$, the limiting renormalized measure vector is obtained as*

$$\mu^{\text{ren}}(0) = \mathscr{P}\chi, \qquad (9)$$

where the matrix operator $\mathscr{P} \equiv \lim_{k\to\infty}(1/k)\sum_{j=0}^{k-1} P^j$ is defined in Theorem A.1.

Proof: Based on Definition 1 and Proposition 1, it suffices to show that the following statement is true:

$$\lim_{\theta\to 0^+} \theta[\mathbb{I} - (1-\theta)\mathbf{P}]^{-1} = \mathscr{P}. \qquad (10)$$

We define the following identities:

$$\Psi(\theta) = \theta[\mathbb{I} - (1-\theta)\mathbf{P}]^{-1} \quad \forall\, \theta \in (0,1) \qquad (11a)$$

$$\Phi(\theta) = \theta\sum_{k=0}^{\infty}(1-\theta)^k\mathscr{P} \quad \forall\, \theta \in (0,1). \qquad (11b)$$

We observe that, for $\theta \in (0,1)$, $\Psi(\theta) = \theta\sum_{k=0}^{\infty}(1-\theta)^k\mathbf{P}^k$. Therefore, $\Psi(\theta) = \Phi(\theta) = \theta\sum_{k=0}^{\infty}(1-\theta)^k(P^k - \mathscr{P})$. Using the result $(P^k - \mathscr{P}) = (P - \mathscr{P})^k$ from Lemma A.1 and Proposition A.1, we have

$$\Psi(\theta) - \Phi(\theta) = \theta[\mathbb{I} - (1-\theta)(P - \mathscr{P})]^{-1}$$
$$\Rightarrow \lim_{\theta\to 0^+}(\Psi(\theta) - \Phi(\theta)) = \lim_{\theta\to 0^+}\theta[\mathbb{I} - (1-\theta)(P - \mathscr{P})]^{-1}.$$

Recalling that for, continuous functions $f(\cdot)$ and $g(\cdot)$,

$$\lim_{\theta\to 0^+} f(\theta) = 0 \text{ and } \lim_{\theta\to 0^+} g(\theta) < \infty$$
$$\Rightarrow \lim_{\theta\to 0^+} f(\theta)g(\theta) = 0,$$

we obtain $(\Psi(\theta) - \Phi(\theta)) = 0 \Rightarrow \lim_{\theta\to 0^+}\Psi(\theta) = \lim_{\theta\to 0^+}\Phi(\theta)$

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

The proof is thus complete. $\qquad\square$

Corollary 1 (to Proposition 2): *The expression $\mathscr{P}\mu^{\text{ren}}(\theta)$ is independent of θ. Specifically, the following identity holds for all $\theta \in (0,1)$:*

$$\mathscr{P}\mu^{\text{ren}}(\theta) = \mathscr{P}\chi. \qquad (12)$$

Proof: Since $[I - (1-\theta)\mathbf{P}]$ is invertible for all $\theta \in (0,1)$, it follow from equation (8) and Theorem A.1 that

$$\mathscr{P}\mu^{\text{ren}}(\theta) = \theta\mathscr{P}\sum_{k=0}^{\infty}(1-\theta)^k\mathbf{P}^k\chi = \theta\sum_{k=0}^{\infty}(1-\theta)^k\mathscr{P}\chi = \mathscr{P}\chi$$

$$\square$$

In the sequel $\mu^{\text{ren}}(0)$ is referred to as the renormalized measure vector unless otherwise stated.

2.1 *Renormalized measure for primitive matrices*

This section presents a special case, where the state transition matrix $\mathbf{P}$ is restricted to be primitive, i.e., irreducible and aperiodic (Berman and Plemmons 1979) (see Definition A.1). In that case, by Perron–Frobenius Theorem (Bapat and Raghavan 1997), $\mathbf{P}$ has exactly one unity eigenvalue and the remaining eigenvalues are located within the unit circle with its center at the origin. Following equation (4), the operator $\mathbf{P}$ is parametrized in terms of $\theta \in (0,1)$ as $\Pi(\theta) \equiv (1-\theta)\mathbf{P}$. Then, $\|\Pi(\theta)\|_\infty = (1-\theta)$ and the eigenvalues of $\Pi(\theta)$ are

$$\lambda_0 = (1-\theta) \text{ and } |\lambda_j| < (1-\theta), \quad j = 1,2,\ldots,(n-1).$$

In this case, Proposition 2 can be restated in a slightly different way as follows.

Proposition 3: *Let p^T be the ℓ_1-normalized left eigenvector of $\mathbf{P}$ corresponding to the unique unity eigenvalue and let $\mathbf{e} \equiv [1 \quad 1 \quad \ldots \quad 1]^T \in \mathbb{R}^n$. The renormalized measure vector is given by the following expression:*

$$\mu^{\text{ren}}(0) = \mathbf{e}\mu^{\text{ren}}, \qquad (13)$$

where the scalar renormalized measure μ^{ren} is denoted as

$$\mu^{\text{ren}} = p^T\chi. \qquad (14)$$

Similarly, Corollary 1 can be restated as Corollary 2.

Corollary 2 (to Proposition 3): *The expression $p^T\mu^{\text{ren}}(\theta)$ is independent of θ. Specifically, the following identity holds for all $\theta \in (0,1)$:*

$$p^T\mu^{\text{ren}}(\theta) = p^T\chi. \qquad (15)$$

Remarks 2: A consequence of Proposition 3 is that the scalar-valued renormalized measure μ^{ren} of the language represents the long-term behavior of plant dynamics in terms of the steady-state state probability vector p and the (assigned) state weight vector χ. Therefore, μ^{ren} in equation (14) is the expected value of plant performance in terms of the assigned weights and probability distribution of the DFSA states.

Remarks 3: Following Corollary 2 the scalar-valued measure μ^{ren} can be readily used as a performance cost

functional for synthesis of optimal discrete-event supervisory control laws under complete observability (Ray *et al.* 2004). This issue is further discussed at the beginning of §3.

Remark 4: The results derived in the finite-dimensional space $\mathbb{R}^n$ for regular languages could be extended to an infinite-dimensional Banach space V for non-regular languages, where the square operator $\mathbf{P}$ is linear and continuous and satisfies the Perron–Frobenius properties analogous to those of a primitive stochastic matrix (Berman and Plemmons 1979). In that case, the sequence of contraction operators $\{(1-\theta_k)\mathbf{P}\}$, with $\theta_k \in (0,1)$ yields a sequence $\mu^{\mathrm{ren}}(\theta_k)$ of bounded vectors converging to the renormalized measure vector $\mu^{\mathrm{ren}}(0) \equiv \mathbf{e}\mu^{\mathrm{ren}}$ as $\theta_k \to 0^+$ with $\mathbf{e}$ being the vector of 1s and the scalar μ^{ren} defined as

$$\mu^{\mathrm{ren}} = \langle p, \chi \rangle, \tag{16}$$

where the vector $p \in \ell_1$ is the unique fixed point of the Perron–Frobenius operator $\mathbf{P}: V \to V$; the bounded linear functional $\chi \in \ell_\infty$ represents (at most countably infinite) state weights; and the expression $\langle u, v \rangle$ is the scalar value of the functional $v \in \ell_\infty$ at the point $u \in \ell_1$. (Note that ℓ_∞ is isometrically isomorphic to the dual space of ℓ_1 (Naylor and sell 1982)).

3. Bounds on derivatives

This section establishes bounds on the derivatives of the renormalized measure $\mu^{\mathrm{ren}}(\theta)$ for all $\theta \in (0,1)$, and computes the limits of the derivatives as $\theta \to 0^+$. This is necessary for developing synthesis tools for optimal discrete-event supervisory control, as a refinement of what is reported in recent literature (Ray *et al.* 2004). By maximizing the renormalized measure $\mu^{\mathrm{ren}}(\theta)$ for a given $\theta \in (0,1)$ an optimal control law can be derived based on the stochastic matrix $\mathbf{P}$ of the supervised plant language and the originally assigned χ-vector. Such an optimal control law can be possibly made θ-independent in the sense of equations (12) and (15) provided that θ is restricted to be not too far away from 0^+. On the other hand, from the perspective of numerical stability and computational accuracy, it is desirable to have a relatively large positive value of θ. The results derived in this section serve toward establishing upper bounds on θ for which the optimal control law should be θ-independent.

A reported algorithm for synthesis of supervisors, which maximizes the language measure (Ray *et al.* 2004), is applicable only for sub-stochastic state

transition cost matrices (i.e., for $\theta > 0$). In this approach, the computational advantage (e.g., convergence in at most n steps, where n is the number of states) offers incentive to investigate the plausibility of dealing with stochastic matrices to take advantage of their analytical and computational properties. For example, identification of explicit bounds on the derivatives of $\mu^{\mathrm{ren}}(\theta)$ provides information on how fast $\mu^{\mathrm{ren}}(\theta)$ can change in the vicinity of 0^+; this approach appears to be far more efficient than numerical search through $\sim 2^{mn}$ possibilities (m being the number of events), which might be infeasible even in relatively simpler systems.

The above observation suggests the possibility of deriving a stochastic optimal policy as the limit of a sequence of sub-stochastic optimal policies, where, optimality is realized in the sense of the measure defined for sub-stochastic transition matrices. The discrete nature of the optimal decision set implies that it actually coincides with some substochastic optimal policy obtained at θ not too close to 0. Note that the explicit bounds on the derivatives of $\mu^{\mathrm{ren}}(\theta)$ would allow θ to be sufficiently removed from 0 so as to be numerically stable. The theoretical work on synthesis of an optimal control law, based on the renormalized measure, largely follows the concept described in the earlier work (Ray *et al.* 2004). This is a topic of current research and will be reported in a forthcoming publication.

Let us start with $\Psi(\theta) \equiv \theta[\mathbb{I} - 1 - \theta)\mathbf{P}]^{-1}$ as defined in equation (11a). The following lemmas are necessary to prove the main propositions in this section.

Lemma 1: $\Psi(\theta)$ *is a non-negative stochastic matrix for all* $\theta \in (0,1]$.

Proof: Let $e \equiv [1 \ 1 \ \cdots \ 1]^T$ be a vector that is compatible for post multiplication with the square matrix $\mathbf{P}$. Since $\mathbf{P}$ is a stochastic matrix, non-negativity of $\Psi(\theta)$ follows from the series expansion

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

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

which implies that $\Psi(\theta)$ is a stochastic matrix as well. $\qquad\square$

Remark 5: The following properties of $\Psi(\theta)$ hold for $\theta \in (0,1)$ by virtue of Theorem A.1.

$$\Psi(\theta)\mathbf{P} = \mathbf{P}\Psi(\theta) \tag{17a}$$

$$\Psi(\theta)\mathscr{P} = \mathscr{P}\Psi(\theta) = \mathscr{P}. \tag{17b}$$

Remark 6: The following relation holds for $k \in \mathbb{N} \cup \{0\}$

$$\frac{\partial^k \mu(\theta)}{\partial \theta^k} = \frac{\partial^k \Psi(\theta)}{\partial \theta^k} \chi. \tag{18}$$

Lemma 2: *For* $\theta \in (0, 1)$

(i) $\quad \dfrac{\partial^k \Psi(\theta)}{\partial \theta^k} = -\left(\dfrac{k}{\theta}\right) \dfrac{\partial^{k-1} \Psi(\theta)}{\partial \theta^{k-1}} \Psi(\theta)\mathbf{P} \quad \forall\, k \in \mathbb{N}\backslash[1]$ (19)

(ii) $\quad \theta^k \dfrac{\partial^k \Psi(\theta)}{\partial \theta^k} = (-1)^k k! \Psi^k(\theta)\mathbf{P}^{k-1}[\Psi(\theta)\mathbf{P} - \mathbb{I}] \quad \forall\, k \in \mathbb{N}.$

$$\tag{20}$$

Proof: Let us derive the first and second derivatives of Ψ and then use the method of induction. An explicit expression for $\partial \mu(\Psi)/\partial \theta$ is obtained as follows:

$$\Psi(\theta) = \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}$$

$$\Rightarrow \Psi(\theta)[\mathbb{I} - (1 - \theta)\mathbf{P}] = \theta$$

$$\Rightarrow \frac{\partial \Psi(\theta)}{\partial \theta}[\mathbb{I} - (1 - \theta)\mathbf{P}] + \Psi(\theta)\mathbf{P} = \mathbb{I}$$

$$\Rightarrow \frac{\partial \Psi(\theta)}{\partial \theta} = [\mathbb{I} - \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}\mathbf{P}]$$

$$[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}. \tag{21}$$

Since $\Psi(\theta)$ and $\mathbf{P}$ commute (see Remark 5), we have

$$\theta \frac{\partial \Psi(\theta)}{\partial \theta} = (-1)\Psi(\theta)(\Psi(\theta)\mathbf{P} - \mathbb{I}). \tag{22}$$

Similarly, the second derivative is explicitly obtained as

$$\theta \frac{\partial \Psi(\theta)}{\partial \theta} = \Psi(\theta) - \Psi(\theta)\Psi(\theta)\mathbf{P}$$

$$\Rightarrow \frac{\partial \Psi(\theta)}{\partial \theta} + \theta \frac{\partial^2 \Psi(\theta)}{\partial \theta^2} = \frac{\partial \Psi(\theta)}{\partial \theta} - \frac{\partial \Psi(\theta)}{\partial \theta}\mathbf{P}\Psi(\theta) - \Psi(\theta)\mathbf{P}\frac{\partial \Psi(\theta)}{\partial \theta}$$

$$\Rightarrow \frac{\partial^2 \Psi(\theta)}{\partial \theta^2} = -2\frac{\partial \Psi(\theta)}{\partial \theta}\Psi(\theta)\mathbf{P}. \tag{23}$$

Commutativity of $\Psi(\theta)$ and $\partial \Psi(\theta)/\partial \theta$ in the last step follows from Remark 5. Hence

$$\theta^2 \frac{\partial^2 \Psi(\theta)}{\partial \theta^2} = (-1)^2 2! \Psi^2(\theta)\mathbf{P}^2(\Psi(\theta)\mathbf{P} - \mathbb{I}). \tag{24}$$

The proof is completed by induction. $\qquad\square$

Corollary 3 (Corollary to Lemma 2)**:**

$$\mathscr{P}\frac{\partial^k \Psi(\theta)}{\partial \theta^k} = \frac{\partial^k \Psi(\theta)}{\partial \theta^k}\mathscr{P} = 0 \quad \forall\, k \in \mathbb{N}\backslash\{0\}. \tag{25}$$

Proof: It follows from Remark 5 and Lemma 2 that $\mathscr{P}\Psi(\theta) = \Psi(\theta)\mathscr{P}$ and $\mathscr{P}\theta^k(\partial^k \Psi(\theta)/\partial \theta^k) = 0\ \forall\, k \in \mathbb{N}/[0]$. The proof follows by setting $\theta \neq 0$. $\qquad\square$

Lemma 3: *For all* $\alpha \neq 0$ *and* $\theta \in (0, 1)$

(i) $\quad \dfrac{\partial \Psi(\theta)}{\partial \theta} = [\mathbb{I} - \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}\mathbf{P}]^2[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}$

$$\tag{26}$$

(ii) $\quad \dfrac{\partial^2 \Psi(\theta)}{\partial \theta^2} = -2\mathbf{P}[\mathbb{I} - \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}\mathbf{P}]^3[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-2}.$

$$\tag{27}$$

Proof: It follows from equation (21) that

$$\frac{\partial \Psi(\theta)}{\partial \theta} = [\mathbb{I} - \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}\mathbf{P}][\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}$$

$$\Rightarrow \frac{\partial \Psi(\theta)}{\partial \theta}[\mathbb{I} - (1 - \theta)\mathbf{P}] = [\mathbb{I} - \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}\mathbf{P}]$$

$$\Rightarrow \frac{\partial \Psi(\theta)}{\partial \theta}[\mathbb{I} - \mathbf{P}] = [\mathbb{I} - \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}\mathbf{P}$$

$$- \theta\frac{\partial \Psi(\theta)}{\partial \theta}\mathbf{P}]$$

$$\Rightarrow \frac{\partial \Psi(\theta)}{\partial \theta}[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}] = [\mathbb{I} - \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}\mathbf{P}$$

$$- \theta\frac{\partial \Psi(\theta)}{\partial \theta}\mathbf{P}]$$

$$\left(\text{by the fact that } \frac{\partial \Psi(\theta)}{\partial \theta}\mathscr{P} = 0 \text{ in equation (25)}\right)$$

$$\Rightarrow \frac{\partial \Psi(\theta)}{\partial \theta} = \left[\mathbb{I} - \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}\mathbf{P} - \theta\frac{\partial \Psi(\theta)}{\partial \theta}\mathbf{P}\right]$$

$$\times [\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}$$

$$\Rightarrow \frac{\partial \Psi(\theta)}{\partial \theta} = \left[\mathbb{I} - \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}\mathbf{P}\right.$$

$$- \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}\mathbf{P}$$

$$\left. + \theta^2[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-2}\mathbf{P}^2\right][\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}$$

$$\Rightarrow \frac{\partial \Psi(\theta)}{\partial \theta} = [\mathbb{I} - \theta[\mathbb{I} - (1 - \theta)\mathbf{P}]^{-1}\mathbf{P}]^2[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}. \tag{28}$$

To prove part (ii) of the lemma, equation (23) is rewritten as

$$\theta\frac{\partial^2\Psi(\theta)}{\partial\theta^2}=-2\frac{\partial\Psi(\theta)}{\partial\theta}\Psi(\theta)\mathbf{P}$$

$$\Rightarrow\frac{\partial^2\Psi(\theta)}{\partial\theta^2}=-2\frac{\partial\Psi(\theta)}{\partial\theta}\mathbf{P}[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1}$$

$$\Rightarrow\frac{\partial^2\Psi(\theta)}{\partial\theta^2}[\mathbb{I}-(1-\theta)\mathbf{P}]=-2\frac{\partial\Psi(\theta)}{\partial\theta}\mathbf{P}$$

$$\Rightarrow\frac{\partial^2\Psi(\theta)}{\partial\theta^2}[\mathbb{I}-\mathbf{P}]=-2\frac{\partial\Psi(\theta)}{\partial\theta}\mathbf{P}+\theta\frac{\partial^2\Psi(\theta)}{\partial\theta^2}\mathbf{P}$$

$$\Rightarrow\frac{\partial^2\Psi(\theta)}{\partial\theta^2}[\mathbb{I}-\mathbf{P}]=2[\mathbb{I}-\theta[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1}\mathbf{P}]^2$$
$$\times[\mathbb{I}-\mathbf{P}+\alpha\mathscr{P}]^{-1}\mathbf{P}[\theta[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1}\mathbf{P}-\mathbb{I}]$$

$$\Rightarrow\frac{\partial^2\Psi(\theta)}{\partial\theta^2}[\mathbb{I}-\mathbf{P}+\alpha\mathscr{P}]=-2\mathbf{P}[\mathbb{I}-\theta[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1}\mathbf{P}]^3$$
$$\times[\mathbb{I}-\mathbf{P}+\alpha\mathscr{P}]^{-1}\left(\text{Since }\frac{\partial^2\Psi(\theta)}{\partial\theta^2}\mathscr{P}=0\text{ in equation (25)}\right)$$

$$\Rightarrow\frac{\partial^2\Psi(\theta)}{\partial\theta^2}=-2\mathbf{P}[\mathbb{I}-\theta[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1}\mathbf{P}]^3[\mathbb{I}-\mathbf{P}+\alpha\mathscr{P}]^{-2}.$$

Corollary 4 (Corollary 1 to Lemma 3):

$$\lim_{\theta\to0^+}\frac{\partial\Psi(\theta)}{\partial\theta}=[\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}-\mathscr{P}.\qquad(29)$$

Proof: It follows from Lemma 3 that $\forall\alpha\neq0$

$$\lim_{\theta\to0^+}\frac{\partial\Psi(\theta)}{\partial\theta}=\lim_{\theta\to0^+}[\mathbb{I}-\theta[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1}\mathbf{P}]^2[\mathbb{I}-\mathbf{P}+\alpha\mathscr{P}]^{-1}$$
$$=[\mathbb{I}-\mathscr{P}\mathbf{P}]^2[\mathbb{I}-\mathbf{P}+\alpha\mathscr{P}]^{-1}$$
$$=[\mathbb{I}-\mathscr{P}]\left\{[\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}+\left(\frac{1}{\alpha}-1\right)\mathscr{P}\right\}$$
$$=[\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}-\mathscr{P}\text{ from Proposition A.2.}$$
$$\square$$

Corollary 5 (Corollary 2 to Lemma 3):

$$\lim_{\theta\to0^+}\frac{\partial\Psi(\theta)}{\partial\theta}=\lim_{\alpha\to\infty}[\mathbb{I}-\mathbf{P}+\alpha\mathscr{P}]^{-1}.\qquad(30)$$

Proof: Proof follows from Corollary 4 and Corollary A.1. $\square$

Corollary 6 (Corollary 3 to Lemma 3):

$$\lim_{\theta\to0^+}\frac{\partial^2\Psi(\theta)}{\partial\theta^2}=2[\mathscr{P}-\mathbf{P}][\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-2}.\qquad(31)$$

Proof: The proof follows from equation (10) and by setting $\alpha=1$ in Proposition A.2. The derivation is similar to Corollary 4. $\square$

The main result on boundedness of the derivatives of $\Psi(\theta)$ (and hence $\mu^{\mathrm{ren}}(\theta)$) are presented as the following two propositions. Specifically, the results in Corollaries 4 and 6 are combined as the following proposition.

Proposition 4:

(i) $\forall k\in\mathbb{N}\setminus\{1\}$
$$\lim_{\theta\to0^+}\frac{\partial^k\Psi(\theta)}{\partial\theta^k}=-k\lim_{\theta\to0^+}\frac{\partial^{k-1}\Psi(\theta)}{\partial\theta^{k-1}}[\mathbf{P}+\mathscr{P}][\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}$$

(ii) $\lim_{\theta\to0^+}\dfrac{\partial^k\Psi(\theta)}{\partial\theta^k}$
$$=\begin{cases}[\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}-\mathscr{P}, & \text{if }k=1\\ (-1)^k k![\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}\\ \quad[\mathbb{I}-[\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}]^{k-1}, & \text{if }k\in\mathbb{N}\setminus\{1\}.\end{cases}$$

Proof: It is observed from Lemma 2 that $\forall k\in\mathbb{N}\setminus\{1\}$ and $\forall\theta\in(0,1)$

$$\theta\frac{\partial^k\Psi(\theta)}{\partial\theta^k}=-k\frac{\partial^{k-1}\Psi(\theta)}{\partial\theta^{k-1}}\Psi(\theta)\mathbf{P}$$
$$\Rightarrow\frac{\partial^k\Psi(\theta)}{\partial\theta^k}=-k\frac{\partial^{k-1}\Psi(\theta)}{\partial\theta^{k-1}}\mathbf{P}[\mathbb{I}-(1-\theta)\mathbf{P}]^{-1}$$
$$\Rightarrow\frac{\partial^k\Psi(\theta)}{\partial\theta^k}[\mathbb{I}-\mathbf{P}]=-k\frac{\partial^{k-1}\Psi(\theta)}{\partial\theta^{k-1}}\mathbf{P}+\theta\frac{\partial^k\Psi(\theta)}{\partial\theta^k}\mathbf{P}$$
$$\Rightarrow\frac{\partial^k\Psi(\theta)}{\partial\theta^k}[\mathbb{I}-\mathbf{P}]=-k\frac{\partial^{k-1}\Psi(\theta)}{\partial\theta^{k-1}}\mathbf{P}-k\frac{\partial^{k-1}\Psi(\theta)}{\partial\theta^{k-1}}\mathbf{P}^2\Psi(\theta)$$
$$\Rightarrow\frac{\partial^k\Psi(\theta)}{\partial\theta^k}=-k\frac{\partial^{k-1}\Psi(\theta)}{\partial\theta^{k-1}}[\mathbf{P}+\Psi(\theta)\mathbf{P}][\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}$$
$$(32)$$
$$\Rightarrow\lim_{\theta\to0^+}\frac{\partial^k\Psi(\theta)}{\partial\theta^k}=-k\lim_{\theta\to0^+}\frac{\partial^{k-1}\Psi(\theta)}{\partial\theta^{k-1}}$$
$$\times\lim_{\theta\to0^+}[\mathbf{P}+\Psi(\theta)\mathbf{P}][\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}$$
$$\Rightarrow\lim_{\theta\to0^+}\frac{\partial^k\Psi(\theta)}{\partial\theta^k}=-k\lim_{\theta\to0^+}\frac{\partial^{k-1}\Psi(\theta)}{\partial\theta^{k-1}}[\mathbf{P}+\mathscr{P}][\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}.$$
$$(33)$$

Part (ii) follows from Corollary 4 for $k=1$ by. For $k\geq2$, it is first noted that

$$[\mathbf{P}+\mathscr{P}][\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}=[\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}-\mathbb{I}+\mathscr{P}$$

It is also noted from Corollary 4 that

$$\lim_{\theta\to0^+}\frac{\partial^k\Psi(\theta)}{\partial\theta^k}=-2\big([\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-2}-[\mathbb{I}-\mathbf{P}+\mathscr{P}]^{-1}\big).$$

Then, the proof follows by induction. $\square$

Remark 5: Note that equation (32) is written for $\alpha=1$ but it is true for any $\alpha\neq0$. Specifically, $\partial^k\Psi(\theta)/\partial\theta^k=-k(\partial^{k-1}\Psi(\theta)/\partial\theta^{k-1})[\mathbf{P}+\Psi(\theta)\mathbf{P}][\mathbb{I}-\mathbf{P}-\alpha\mathscr{P}]^{-1}$ for all $\alpha\neq0$.

The next proposition establishes bounds on the derivatives of $\mu(\theta)$ in an elementwise sense by computing bounds on the induced sup-norm of the derivatives of $\Psi(\theta)$. Recall that χ has been defined to have unity infinity norm.

Proposition 5:

$$\left\|\frac{\partial^k \mu(\theta)}{\partial \theta^k}\right\|_\infty \leq k! 2^{k+1} \left(\inf_{a \neq 0}\left\|[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\right\|_\infty\right)^k \quad \forall\, \theta \in [0, 1].$$

Proof: We recall that $[\mathbb{I} - \mathbf{P} - \alpha\mathscr{P}]^{-1}$ is well defined for all $\alpha \neq 0$. (Proposition A.1) which implies

$$\inf_{a \neq 0}\left\|[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\right\|_\infty < \infty. \tag{34}$$

The assertion follows by taking induced sup-norm on equation (28) with right multiplication by χ and taking infimum.

$$\left\|\frac{\partial \mu(\theta)}{\partial \theta}\right\|_\infty = \left\|\frac{\partial \Psi(\theta)}{\partial \theta}\chi\right\|_\infty$$
$$\leq \left\|[\mathbb{I} - \theta[\mathbb{I} - (1-\theta)\mathbf{P}]^{-1}\mathbf{P}]^2[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\right\|_\infty \tag{35}$$
$$\leq \left\|[\mathbb{I} - \theta[\mathbb{I} - (1-\theta)\mathbf{P}]^{-1}\mathbf{P}]^2\right\|_\infty \left\|[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\right\|_\infty$$
$$\leq (1+1)^2\left\|[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\right\|_\infty. \tag{36}$$

It follows from equation (32) and Remark 5 that

$$\left\|\frac{\partial^k \mu(\theta)}{\partial \theta^k}\right\|_\infty = \left\|\frac{\partial^k \Psi(\theta)}{\partial \theta^k}\chi\right\|_\infty$$
$$= \left\|-k\frac{\partial^{k-1}\Psi(\theta)}{\partial \theta^{k-1}}[\mathbf{P} + \Psi(\theta)\mathbf{P}][\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\chi\right\|_\infty$$
$$\leq k\left\|\frac{\partial^{k-1}\Psi(\theta)}{\partial \theta^{k-1}}\right\|_\infty (1+1)\left\|[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\right\|_\infty$$
$$\leq 2k\left\|\frac{\partial^{k-1}\Psi(\theta)}{\partial \theta^{k-1}}\right\|_\infty \left\|[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\right\|_\infty \tag{37}$$

The proof follws by induction. $\square$

Remark 6: Note that for the second derivative,

$$\left\|\frac{\partial^2 \mu(\theta)}{\partial \theta^2}\right\|_\infty \leq 16\left(\inf_{a \neq 0}\left\|[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\right\|_\infty\right)^2 \quad \forall\, \theta \in [0, 1].$$
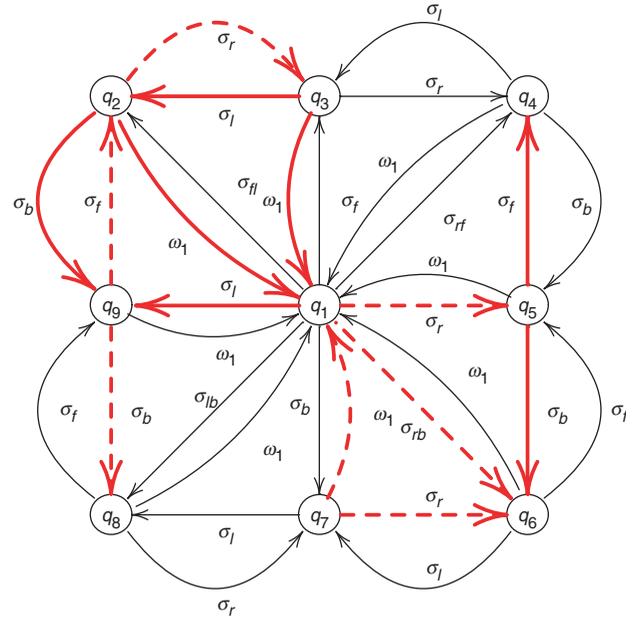
Figure 1. DFSA Models of the (unsupervised and supervised) plant: red thick dashed-line arcs are controllable transitions that are enabled for the unsupervised plant and are later disabled under the control action; red thick solid-line arcs are controllabed transitions that remain enabled under the control action; black thin solid-line arcs are uncontrollable transitions.

4. An illustrative example

This section presents a simple example to illustrate the underlying concept of renormalized measure and its usage for synthesis of discrete-event supervisory controllers (Ray *et al.* 2004). Let us consider a deterministic finite state automaton (DFSA) with the state set $Q = \{q_1, \ldots, q_9\}$ and the event alphabet $\Sigma = \{\sigma_r, \sigma_i, \sigma_f, \sigma_b, \sigma_{fi}, \sigma_{rf}, \sigma_{rb}, \sigma_{ib}, \omega_1\}$. The DFSA model of the unsupervised plant is shown in figure 1. The arcs, shown in thick red lines, are controllable and those shown in thin black lines are uncontrollable. (Note that controllability is state-dependent, i.e., the same event could be controllable at one state and uncontrollable at another state.) The arcs with thick red dashed lines in figure 1 indicate that the event(s) on this arc have been disabled under the action of a supervisory controller. Table 1 lists the stochastic transition martix $\mathbf{P}$ which is primitive because $\mathbf{P}^2$ is a positive matrix. The stationary state probability vector (i.e., the left eigen-vector p^T corresponding to the unique unity eigenvalue) of $\mathbf{P}$ is

$$p^T = [\,0.234 \quad 0.041 \quad 0.065 \quad 0.084 \quad 0.095 \quad 0.016 \quad 0.153 \quad 0.147 \quad 0.076\,]$$

Table 1. Stochastic transtition matrix **P**.

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
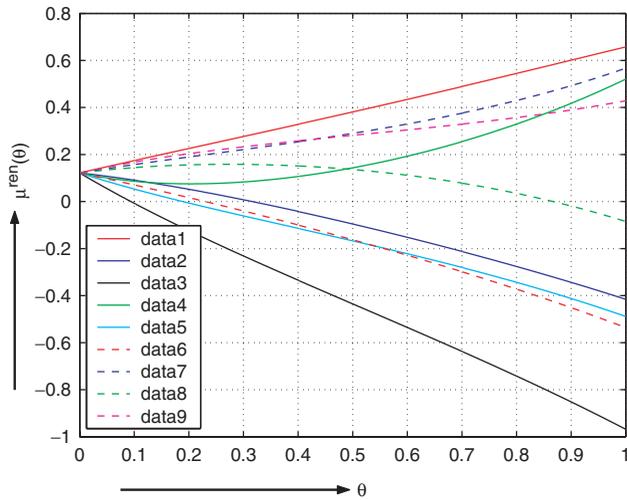
Figure 2. Variations of $\mu^{\text{ren}}(\theta)$ for the unsupervised plant.
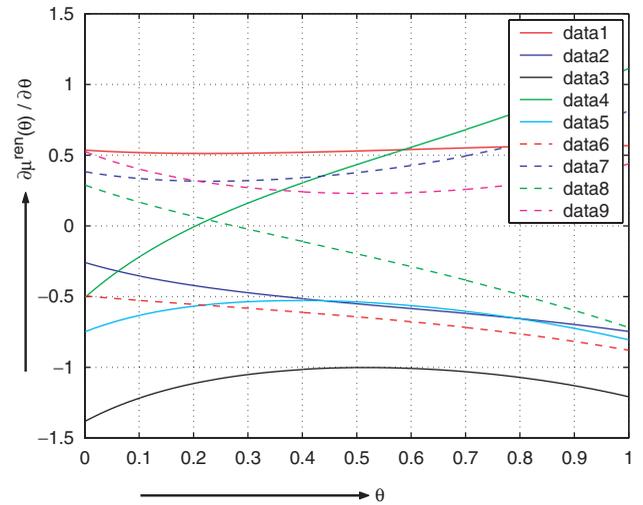
Figure 3. Slope $\mu^{\text{ren}}(\theta)$ for unsupervised plant.

The χ-vector for the DFSA is chosen as

$$\chi = [\,0.66 \quad -0.42 \quad -0.97 \quad 0.52 \quad -0.49 \quad -0.57 \quad 0.57 \quad -0.09 \quad 0.43\,]^T.$$

Variations of the elements of the renormalized measure vector $\mu^{\text{ren}}(\theta)$ with θ are plotted in figure 2, where all elements of $\mu^{\text{ren}}(\theta)$ converage to a single value $\mu^{\text{ren}} \approx 0.122$ as $\theta \to 0^+$

Figure 3 and 4 respectively show the variations in the gradient (i.e., first derivative) and the curvature (i.e., second derivative) of the renormalized measure vector $\mu^{\text{ren}}(\theta)$ with θ. An optimal control policy (Ray *et al*. 2004), which maximizes the renormalized measure as $\theta \to 0^+$, is computed by numerically searching for the automation configuration. Figure 1 shows the optimally supervised plant, where the red thick dashed-line arcs show disabled controllable transitions.

Figure 5 displays a comparison of the histograms of the state probability vector p for the unsupervised and optimally supervised plant; the respective values of $\mu^{\text{ren}} = p^T \chi$ in equation (14) are 0.1218 and 0.3534, which indicates that the long term performance is almost tripled under optimal supervision. The implication is that the probability of terminating in favorable states has increased and that of being in the unfavorable states has decreased.

5. Summary and conclusion

This paper extends the notion of normalized signed measure of probabilistic regular languages (Ray 2005) through renormalization with respect to the measure of the maximal language Σ^* over the given alphabet Σ. The work reported in this paper removes a fundamental
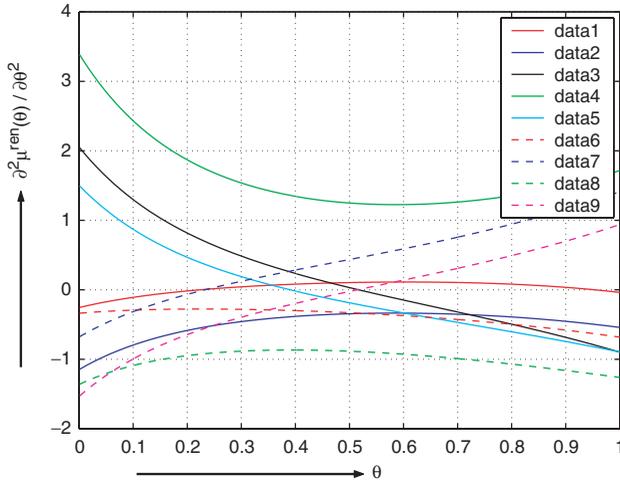
Figure 4. Curvature of $\mu^{\mathrm{ren}}(\theta)$ for the unsupervised plant.
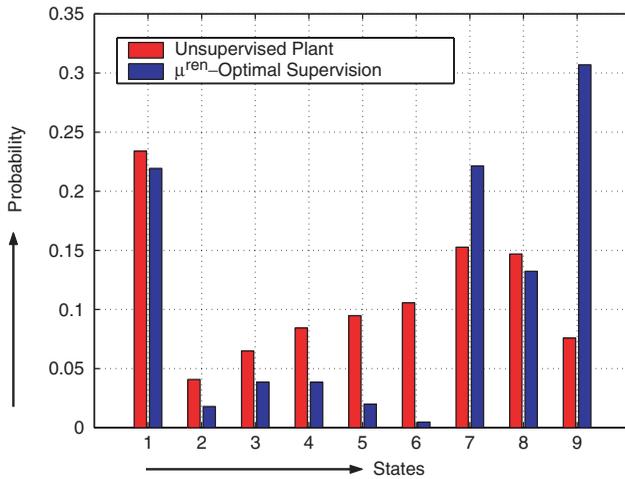
Figure 5. Probability distributions over the automation states.

restriction of earlier analysis, namely, each row sum of the state transition cost matrix being strictly less than 1, instead of being equal to 1 (Ray 2005, Ray *et al.* 2005). It also shows that if the (primitive stochastic) state transition matrix has a unique unity eigenvalue and the magnitude of each of the remaining eigenvalues is strictly less than unity, then the scalar-valued renormalized measure μ^{ren} represents the expected value of the plant performance in terms of the assigned weights and probability distribution of the plant states. The concept is illustrated for optimal supervisory control synthesis by an example. Future research areas include:

- Optimal control of regular languages under complete and partial observability with the scalar μ^{ren} as the cost functional.
- Renormalized measure of non-regular languages.

Appendix

Properties of Stochastic Matrices

This appendix presents pertinent results for stochastic matrices that have been used to derive the properties of the renormalized measure in §§ 2 and 3. Let P be the state transition probability matrix of a stationary Markov chain with finitely many states. Then, P is a stochastic (i.e., non-negative with each row sum being identically equal to unity (Bapat and Raghavan 1997).

Theorem A.1: *For any stochastic matrix* **P**, *the following limit exists.*

$$\lim_{k \to \infty} \frac{1}{k} \sum_{j=0}^{k-1} \mathbf{P}^j \to \mathscr{P}, \tag{38}$$

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

$$\mathbf{P}\mathscr{P} = \mathscr{P}\mathbf{P} = \mathscr{P} = \mathscr{P}^2. \tag{39}$$

Proof: The proof is given in Bapat and Raghavan (1997, pp. 50–51). □

Definition A.1: *A stochastic matrix* **P** *is irreducible and aperiodic, also called primitive* (Berman and Plemmons 1979), *if there exists $k \in \mathbb{N}$ such that $\mathbf{P}^k > 0$* (Bapat and Raghavan 1997, p. 49).

Theorem A.2: *For any irreducible and aperiodic stochastic matrix* **P**, *the following limit exists.*

$$\lim_{k \to \infty} \mathbf{P}^k \to \mathscr{P}, \tag{40}$$

where $\mathscr{P}$ is given in Definition A.1. *In this case, each row of $\mathscr{P}$ is identically equal to the unique probability vector p^T that is a fixed point of the operator* **P**, *i.e., $p^T \mathbf{P} = p^T$.*

Proof: The proof is given in Bapat and Raghavan (1997, pp. 49–50). □

Remark A.1: Since **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 **P**. However, the next proposition shows that $[\mathbb{I} - \mathbf{P} + \mathscr{P}]$ is always invertible.

Proposition A.1: *Let* **P** *be a stochastic matrix. Then* $[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}$ *is well-defined for all* $\alpha \neq 0$.

Proof: Let $\alpha \neq 0$ be arbitrary. Let us assume that $[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]$ is not invertible. Then, it follows that there exists a vector $\vartheta \neq 0$ such that $[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]\vartheta \neq 0$.

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

$$\Rightarrow [\alpha\mathscr{P} - \alpha^2\mathscr{P}]\vartheta = \alpha\mathscr{P}\vartheta$$

$$\Rightarrow \alpha^2\mathscr{P}\vartheta = 0 \Rightarrow \mathscr{P}\vartheta = 0 \quad \text{since } \alpha \neq 0.$$

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

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

This is a contradiction. (Note that the commutative and idempotent properties of $\mathscr{P}$ in Theorem A.1 have been used). $\qquad\square$

The next proposition is useful for finding the properties of the renormalized measure in § 3 and requires the following lemma.

Lemma A.1:

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

Proof: The above identity is readily seen to be 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 Theorem A.1. The proof of the lemma follows directly by the method of induction. $\qquad\square$

Proposition A.2: *The following identity holds for any stochastic matrix* **P**.

$$[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1} = [\mathbb{I} - \mathbf{P} + \mathscr{P}]^{-1} + \left(\frac{1-\alpha}{\alpha}\right)\mathscr{P} \quad \forall\, \alpha \neq 0. \tag{41}$$

Proof: The matrix $[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]$ is invertible for all $\alpha \neq 0$ by Proposition A.1. By setting $\alpha = 1$, Lemma A.1. yields $[\mathbb{I} - \mathbf{P} + \mathscr{P}]^{-1} = \mathbb{I} + \sum_{k=1}^{\infty}(\mathbf{P}^k - \mathscr{P})$. The proof is completed from the following identity (by use of Theorem A.1):

$$[\mathbb{I} - \mathbf{P} + \mathscr{P}]\left(I + \sum_{k=1}^{\infty}(\mathbf{P}^k - \mathscr{P})\right) + \left(\frac{1-\alpha}{\alpha}\right)$$

$$\times [\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]\mathscr{P} = \mathbb{I}. \qquad\square$$

The following two corollaries of Proposition A.2 are useful for finding bounds on the derivatives of the renormalized measure in § 3.

Corollary A.1 (Corollary 1 to Proposition A.2):

$$\lim_{\alpha\to\infty}[\mathbb{I} - \mathbf{P} + \mathscr{P}]^{-1} = [\mathbb{I} - \mathbf{P} + \mathscr{P}]^{-1} - \mathscr{P} \tag{42}$$

Proof: The proof follows from Proposition A.2. $\qquad\square$

Corollary A.2 (Corollary 2 to Proposition A.2):

(i) $\quad \left\|[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\mathscr{P}\right\|_\infty = \left|\dfrac{1}{\alpha}\right| \quad \forall\, \alpha \neq 0 \tag{43a}$

(ii) $\quad \left\|[\mathbb{I} - \mathscr{P} + \alpha\mathscr{P}]^{-1}\right\|_\infty \geq \left|\dfrac{1}{\alpha}\right| \quad \forall\, \alpha \neq 0 \tag{43b}$

(iii) $\quad \left\|[I - \mathbf{P} + \mathscr{P}]^{-1}\right\|_\infty \geq 1, \tag{43c}$

where $\|\bullet\|$ *indicates the induced sup-norm of the operator* $\bullet$.

Proof: Expressing $[\mathbb{I} - \mathbf{P} + \mathscr{P}]^{-1}$ as a convergent infinite sum following Proposition A.1, we have

$$[\mathbb{I} - \mathbf{P} + \mathscr{P}]^{-1}\mathscr{P} = \sum_{k=0}^{\infty}(\mathbf{P} - \mathscr{P})^k\mathscr{P} = \mathscr{P}.$$

Hence,

$$[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\mathscr{P} = [I - \mathbf{P} + \mathscr{P}]^{-1}\mathscr{P} - \mathscr{P} + \frac{1}{\alpha}\mathscr{P} = \frac{1}{\alpha}\mathscr{P}$$

which implies

$$\left\|[\mathbb{I} - \mathbf{P} + \alpha\mathscr{P}]^{-1}\mathscr{P}\right\|_\infty = \left|\frac{1}{\alpha}\right|\|\mathscr{P}\|_\infty = \left|\frac{1}{\alpha}\right|$$

The last step follows from the stochastic property of $\mathscr{P}$ (see Theorem A.1), which implies that each row sum of $\mathscr{P}$ is 1 and hence $\|\mathscr{P}\|_\infty = 1$. The second assertion follows from the first by using the sub-multiplicative property of induced norms. The third assertion is immediate from the second by setting $\alpha = 1$. $\qquad\square$

References

R.B. Bapat and T.E.S. Raghavan, *Nonnegative Matrices and Application*, Cambridge, UK: Cambridge University Press, 1997.

A. Berman and R.J. Plemmons, *Nonnegative Matrices in the Mathematical Science*, New York: Academic Press, 1979.

V.K. Garg, "Probabilistic languages for modeling of DEDs", in *Proceedings of 1992 IEEE Conference on Information and Sciences*, Princeton, NJ, 1992a, pp. 198–203.

V.K. Garg, "An algebraic approach to modeling probabilistic discrete event systems", in *Proceedings of 1992 IEEE*

Conference on Decision and Control, Tucson, AZ, 1992b, pp. 2348–2353.

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

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

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

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

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