Approximating Stochastic Events by Quantum Automata*

Alberto Bertoni[1 †] Carlo Mereghetti[1 ‡] Beatrice Palano[1 §]

[1] *Dipartimento di Scienze dell'Informazione, Università degli Studi di Milano via Comelico 39/41, 20135 Milano, Italy.*

Abstract. Given a class $\{p_\alpha \mid \alpha \in I\}$ of events induced by M-state 1qfa's on alphabet Σ, we investigate the size (number of states) of 1qfa's δ-approximating convex linear combinations of $\{p_\alpha \mid \alpha \in I\}$. We obtain:

- A $O((M/\delta^3)\, d \log^2(d/\delta^2))$ size bound, where d is the Vapnik dimension of $\{p_\alpha(w) \mid w \in \Sigma^*\}$.
- A $O((M/\delta^2) \log n)$ size bound, for p_α's n-periodic. This shows the existence of a Monte Carlo 1qfa recognizing the language $L_n = \{a^{kn} \mid k \in \mathbf{N}\}$ with bounded error ε and $O((1/\varepsilon^3) \log n)$ states.
- A $O((1/\delta^2) \log n)$ size bound for inducing a δ-approximation of $(1+p)/2$, for any n-periodic event p whose discrete Fourier transform has ℓ_1-norm not exceeding n.

Keywords: stochastic events, quantum automata

1 Introduction

1-way quantum finite automata (1qfa's, for short) [2, 4, 6, 7] are computational devices particularly interesting since they represent a theoretical model for a quantum computer with finite memory.

Formally, a (measure-once [3, 5, 9]) 1qfa with q control states on the input alphabet Σ is a system $A = (\pi, U(\sigma), P)$, where $\pi \in \mathbf{C}^{1\times q}$, for each $\sigma \in \Sigma$, $U(\sigma) \in \mathbf{C}^{q\times q}$ is a *unitary* matrix, and $P \in \mathbf{C}^{q\times q}$ is a projector that biunivocally individuate the observable $\mathcal{O} = 1 \cdot P + 0 \cdot (I - P)$. The stochastic event induced by A is the function $p_A : \Sigma^* \to [0,1]$ defined by $p_A(\sigma_1 \ldots \sigma_k) = \left\| \pi \left(\prod_{i=1}^k U(\sigma_i) \right) P \right\|^2$, with $\|\ \|$ the vector norm.

In this work, we study the size (number of states) of 1qfa's whose induced events approximate given stochastic events in the following sense:

Definition 1 *A δ-approximation in L^∞ of a given stochastic event $p : \Sigma^* \to [0,1]$ is any stochastic event $q : \Sigma^* \to [0,1]$ satisfying* $\sup_{w\in\Sigma^*} \{|p(w) - q(w)|\} \le \delta$.

2 Approximating the convex closure of classes of stochastic events

Given a family $\mathcal{F} = \{\varphi_\alpha : \Sigma^* \to [0,1] \mid \alpha \in I\}$ of stochastic events induced by M-state 1qfa's $(\pi_\alpha, U_\alpha(\sigma), P_\alpha)$, let $\tilde{\mathcal{F}}$ be the convex closure of $\mathcal{F}$, i.e., the class of stochastic events ξ obtained as convex linear combination $\xi(w) = \sum_{\alpha\in I} b_\alpha \varphi_\alpha(w)$ ($b_\alpha \ge 0$, $\sum_{\alpha\in I} b_\alpha = 1$). We are interested in estimating the number of states of 1qfa's inducing events that δ-approximate $\xi \in \tilde{\mathcal{F}}$.

Since $b_\alpha \ge 0$ and $\sum_{\alpha\in I} b_\alpha = 1$, we can interpret b_α's as a probability distribution on I. For any $w \in \Sigma^*$, $\varphi_\alpha(w)$ becomes a random variable with expectation $E[\varphi_\alpha(w)] = \sum_{\alpha\in I} b_\alpha \varphi_\alpha(w) = \xi(w)$. We can approximate such an expectation by an empirical average of the events in $\mathcal{F}$. To this purpose, we design the following algorithm:

for $t := 1$ **to** S **do**
  $\alpha[t] := \alpha$ independently chosen in I
    with probability b_α;
output the 1qfa A defined as

$$A = \left(\sqrt{1/S} \bigoplus_{t=1}^{S} \pi_{\alpha[t]}, \bigoplus_{t=1}^{S} U_{\alpha[t]}(\sigma), \bigoplus_{t=1}^{S} P_{\alpha[t]} \right),$$

where '$\oplus$' denotes matrix direct sum.

First of all, observe that the stochastic event $\psi_S : \Sigma^* \to [0,1]$ induced by A is defined, for any $w \in \Sigma^*$, as $\psi_S(w) = (1/S) \sum_{t=1}^S \varphi_{\alpha[t]}(w)$, i.e., ψ_S is an empirical average of the events in $\mathcal{F}$. Now, if

$$\mathrm{Prob}\left\{ \sup_{w\in\Sigma^*} \{|\xi(w) - \psi_S(w)|\} \ge \delta \right\} < 1 \qquad (1)$$

holds true, then the existence of an $(S \cdot M)$-state 1qfa inducing a δ-approximation of ξ is guaranteed.

Estimating

$$\mathrm{Prob}\left\{ \sup_{w\in\Sigma^*} \left\{ \left| \frac{1}{S} \sum_{t=1}^{S} \varphi_{\alpha[t]}(w) - E[\varphi_\alpha(w)] \right| \right\} \ge \delta \right\}$$

is a classical problem of uniform convergence of empirical averages to their expectations [1, 10].

2.1 General framework

A general solution can be given in terms of the Vapnik dimension of the class of random variables $\{\varphi_\alpha(w) \mid w \in \Sigma^*\}$ (see [1] for more details).

Definition 2 *Given a class $\mathcal{F}$ of functions $\psi : I \to [0,1]$ and $\beta \in (0,1)$, a subset $A \subset I$ is said to be* shattered *by $\mathcal{F}$ if, for every $X \subset A$, there exists $\varphi \in \mathcal{F}$ for which $x \in X$ implies $\varphi(x) \ge \beta$, and $x \in A - X$ implies $\varphi(x) < \beta$. Then the* Vapnik dimension V-dim$\{\mathcal{F}\}$ *is the maximal cardinality of shattered subsets of I.*

---

*Partially supported by MURST, under the projects "Linguaggi formali: teoria ed applicazioni" and "FIRB: Complessità descrizionale di automi e strutture correlate".

† bertoni@dsi.unimi.it

‡ mereghetti@dsi.unimi.it

§ palano@dsi.unimi.it

As an easy consequence of Theorem 3.6 in [1], for $S = O(\frac{1}{\delta^3} d \log^2 \frac{d}{\delta^2})$, where $d = \text{V-dim}\{\varphi_\alpha(w) \mid w \in \Sigma^*\}$, we have $\text{Prob}\left\{\sup_{w\in\Sigma^*}\{|\xi(w) - \psi_S(w)|\} \geq \delta\right\} < 1$. Therefore we can conclude:

Theorem 3 *If $\{\varphi_\alpha(w) \mid w \in \Sigma^*\}$ is a class of stochastic events induced by M-state 1qfa's, then every convex linear combination $\xi(w) = \sum_{\alpha\in I} b_\alpha \varphi_\alpha(w)$ can be δ-approximated by a 1qfa with $O(\frac{M}{\delta^3} d \log^2 \frac{d}{\delta^2})$ states, where $d = \text{V-dim}\{\varphi_\alpha(w) \mid w \in \Sigma^*\}$.*

2.2 The unary periodic case

We directly solve the problem in the very simple case of periodic events. We consider the class $\mathcal{F} = \{\varphi_\alpha : \{a\}^* \to [0,1] \mid \alpha \in I\}$ where every φ_α is an n-periodic event. Then, we rewrite Equation (1) by considering the union bound and Höffdings' inequality [8] as

$$\text{Prob}\left\{\sup_{0\leq k<n}\{|\psi_S(a^k) - \xi(a^k)|\} \geq \delta\right\} \leq n \cdot 2\mathrm{e}^{-2\delta^2 S}.$$

By requiring $n \cdot 2\mathrm{e}^{-2\delta^2 S} < 1$, we obtain

Theorem 4 *Given a family Ψ of n-periodic events induced by M-state 1qfa's, any event in the convex closure of Ψ can be δ-approximated by the event induced by a 1qfa with $O((M/\delta^2)\log n)$ states.*

We can apply this latter result to language recognition. A *unary language* $L \subset \{a\}^*$ is said to be recognized by a 1qfa A in *Monte Carlo mode* if and only if there exists $\varepsilon \in (0, 1/2)$ such that, for any $k \geq 0$: $a^k \in L$ implies $p_A(a^k) = 1$, $a^k \notin L$ implies $p_A(a^k) \leq \varepsilon$.

Consider the language $L_n = \{a^{kn} \mid k \in \mathbf{N}\}$. We get, thus improving [2]

Theorem 5 *For any $n > 1$, there exists a 1qfa accepting L_n in Monte Carlo mode with bounded error ε and $O((1/\varepsilon^3)\log n)$ states.*

3 Approximating a family of periodic events

We present a class of n-periodic events that are approximable by events induced by $O(\log n)$-state 1qfa's.

Let $p : \{a\}^* \to [0,1]$ be an n-periodic event completely characterized by the vector $(p(\varepsilon), p(a), \ldots, p(a^{n-1}))$. Its *discrete Fourier transform* is the complex vector $P = (P(0), \ldots, P(n-1))$ such that $P(j) = \sum_{k=0}^{n-1} p(a^k)\mathrm{e}^{i\frac{2\pi}{n}kj}$. For any $k \geq 0$, we have $p(a^k) = \frac{1}{n}\sum_{j=0}^{n-1} P(j)\,\mathrm{e}^{-i\frac{2\pi}{n}kj}$. The ℓ_1-norm of P is $\| P \|_1 = \sum_{j=0}^{n-1} |P(j)|$.

Theorem 6 *Let $p : \{a\}^* \to [0,1]$ be an n-periodic event whose discrete Fourier transform P satisfies $\| P \|_1 = n$. Then, the event $(1+p)/2$ is δ-approximable by the event induced by a 1qfa with $O((1/\delta^2)\log n)$ states.*

Proof. We can expand p by its discrete Fourier transform P. Setting $P(j) = \rho_j \mathrm{e}^{i\vartheta_j}$ and p ranging in $[0,1]$, we get $p(a^k) = \sum_{j=0}^{n-1} \frac{\rho_j}{n} \cos\left(\frac{2\pi}{n}kj - \vartheta_j\right)$. Since $\| P \|_1 = \sum_{j=0}^{n-1} \rho_j = n$, we can interpret ρ_j/n as a probability distribution on $\mathbf{Z}_n$. Any event $\phi_j(a^k) = \cos^2\left(\frac{\pi}{n}kj - \frac{\vartheta_j}{2}\right)$ is induced by a 2-state 1qfa. By applying the algorithm in Section 2, and considering Theorem 4, there exists a 1qfa with $O((1/\delta^2)\log n)$ states inducing the stochastic event $\psi(a^k) = \frac{1}{S}\sum_{t=1}^{S} \cos^2\left(\frac{\pi}{n}kj[t] - \frac{\vartheta_{j[t]}}{2}\right)$, which is a δ-approximation of the event $(1+p)/2$ for $S = O((1/\delta^2)\log n)$. □

This result can be easily extended to encompass periodic events for which $\| P \|_1 \leq n$.

Again, we can apply this result to language recognition. Given an event $p : \{a\}^* \to [0,1]$ and a real $\lambda \in [0,1]$, *the unary language L_λ defined by p with cut-point λ* writes as $L_\lambda = \{a^k \mid k \in \mathbf{N},\ p(a^k) > \lambda\}$. The cut-point is said to be *isolated* if there exists a positive real δ such that $|p(a^k) - \lambda| \geq \delta$, for any $k \geq 0$. Moreover, if p is induced by a 1qfa A then L_λ is said to be recognized by A with cut point λ (isolated by δ).

We can immediately obtain

Theorem 7 *Let $p : \{a\}^* \to [0,1]$ be an n-periodic event whose discrete Fourier transform P satisfies $\| P \|_1 \leq n$, and let L be a unary language defined by p with cut point λ isolated by 2δ. Then L can be recognized by a 1qfa with cut point $\frac{1}{2} + \frac{1}{2}\lambda$ isolated by δ and $O((1/\delta^2)\log n)$ states.*

References

[1] N. Alon, S. Ben-David, N. Cesa-Bianchi and D. Haussler. Scale-Sensitive Dimensions, Uniform Convergence, and Learnability. *J. ACM*, 44:615–631, 1997.

[2] A. Ambainis and R. Freivalds. 1-way quantum finite automata: strengths, weaknesses and generalizations. In *Proc. 39th FOCS*, pp. 332–342, 1998.

[3] A. Bertoni and M. Carpentieri. Regular languages accepted by quantum automata. *Information and Computation*, 165:174–182, 2001.

[4] A. Bertoni, C. Mereghetti and B. Palano. Quantum computing: 1-way quantum automata. In *Proc. 7th Conf. Develop. Lang. Th.*, LNCS 2710, pp. 1–20, Springer, 2003.

[5] A. Brodsky and N. Pippenger. Characterizations of 1-way quantum finite automata. *SIAM J. Comput.*, 31:1456-1478, 2001.

[6] J. Gruska. *Quantum Computing*. McGraw-Hill, 1999.

[7] J. Gruska. Descriptional complexity issues in quantum computing. *J. Aut. Lan. Comb.*, 5:191–218, 2000.

[8] W. Höffdings. Probability inequalities for sums of bounded random variables. *J. American Statistical Association*, 58:13–30, 1963.

[9] C. Moore and J. Crutchfield. Quantum automata and quantum grammars. *Theoretical Computer Science*, 237:275–306, 2000.

[10] V.N. Vapnik and A.Y. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Th. Prob. App.*, 16:264–280, 1971.