Theoretical Computer Science 551 (2014) 102–115

Contents lists available at ScienceDirect

Theoretical Computer Science

www.elsevier.com/locate/tcs

Size lower bounds for quantum automata ☆

Maria Paola Bianchi, Carlo Mereghetti *, Beatrice Palano

Dipartimento di Informatica, Università degli Studi di Milano, via Comelico 39, 20135 Milano, Italy

A R T I C L E   I N F O

Article history:
Received 24 December 2013
Received in revised form 14 April 2014
Accepted 6 July 2014
Available online 15 July 2014
Communicated by M. Hirvensalo

Keywords:
Quantum finite automata
Descriptional complexity

A B S T R A C T

We compare the descriptional power of quantum finite automata with control language (QFCS) and deterministic finite automata (DFAS). By suitably adapting Rabin's technique, we show how to convert any given QFC to an equivalent DFA, incurring in an at most exponential size increase. This enables us to state a lower bound on the size of QFCS, which is logarithmic in the size of equivalent minimal DFAS. In turn, this result yields analogous size lower bounds for several models of quantum finite automata in the literature.

1. Introduction

While we can hardly expect to see a full-featured quantum computer in the near future, it is reasonable to envision classical computing devices incorporating quantum components. Since the physical realization of quantum systems has proved to be a complex task, it is also reasonable to keep quantum components as "small" as possible. Small size quantum devices are modeled by *quantum finite automata* (QFAS), a theoretical model for quantum machines with finite memory. Thus, it is well worth investigating, from a theoretical point of view, *lower limits* to the size of QFAS when performing certain tasks, also emphasizing *trade-offs* with the size of equivalent classical devices.

Originally, two models of QFAS are proposed: *measure-once* QFAS [10,21], where the probability of accepting words is evaluated by "observing" just once, at the end of input processing, and *measure-many* QFAS [3,17], having such an observation performed after each move. Several modifications to these two original models of QFAS, motivated by different possible physical realizations, are then proposed. Thus, e.g., enhanced [22], reversible [12], Latvian [2], and measure-only QFAS [8] are introduced. Results in the literature (see, e.g., [2,5,11,18]) show that all these quantum variants are strictly less powerful than deterministic finite automata (DFAS), although retaining a higher descriptional power (i.e., their sizes can be significantly smaller than their equivalent classical devices).

To enhance the low computational power of these "purely quantum" systems, *hybrid models* featuring both a quantum and a classical component are studied. Examples of such hybrid systems, all reaching the same computational power of classical automata, are QFAS with open time evolution [13], QFAS with quantum and classical states (QCFAS) [28], and QFAS *with control language* (QFCS) [5,19].

Here, we are interested in this latter model which, roughly speaking, can be described as follows. A QFC $\mathcal{A}$ can be regarded to as a computational device having a quantum processor, namely a QFA, controlled by a DFA. The state of the QFA is observed after each move by an observable with a fixed, but arbitrary, set of possible outcomes. On any given

☆ Partially supported by the Italian MIUR under the project PRIN-2010LYA9RH_005 "PRIN: Automi e Linguaggi Formali: Aspetti Matematici e Applicativi." A preliminary version of this work was presented at the *11th Int. Conf. Unconventional Computation & Natural Computation (UCNC 2013), Milano, Italy, July 1–5, 2013.*

* Corresponding author.
E-mail addresses: bianchi@di.unimi.it (M.P. Bianchi), mereghetti@di.unimi.it (C. Mereghetti), palano@di.unimi.it (B. Palano).

http://dx.doi.org/10.1016/j.tcs.2014.07.004
0304-3975/

input word x, a sequence y of outcomes is generated with a certain probability. The computation of $\mathcal{A}$ on x is accepting whenever y belongs to the regular language (the control language) recognized by the DFA. In [5,19], it is proved that the class of languages accepted with isolated cut point by QFCs coincides with the class of regular languages, and that QFCs can be exponentially smaller than their equivalent classical automata. It may be worth quickly noticing that a relevant difference between QFCs and QCFAs [28] is to be pointed out in the communication policy between the two internal components: in QCFAs a two-way information exchange between the classical and quantum parts is established, while in QFCs only the quantum component affects the dynamic of the classical one.

A relevant feature of QFCs, of interest in this paper, is that they can naturally and directly simulate several models of QFAs by preserving the size. This property makes QFCs a general unifying framework within which to investigate size results for different quantum paradigms: size lower bounds or size trade-offs proved for QFCs may directly apply to simulated types of QFAs as well. In fact, the need for a general quantum framework is witnessed by several results in the literature (see, e.g., [1,3,4,6,7,9,20,27]) showing that QFAs can be exponentially more succinct than equivalent classical automata, by means of techniques which are typically targeted on the particular type of QFA and not easily adaptable to other paradigms. So, to cope with this specialization problem, here we study size lower bounds and trade-offs for QFCs.

After introducing some basic notions in Section 2, we show in Section 3 how to build from a given QFC an equivalent DFA. To this aim, we must suitably modify classical Rabin's technique [24] since the equivalence relation we choose to define the state set of the DFA is not a congruence. On the other hand, this relation – based on Euclidean norm – allows us to directly estimate the cost of the conversion by a geometrical argument on compact spaces. We obtain that the size of the resulting DFA is at most exponentially larger than the size of the QFC. Stated in other terms in Section 4, this latter result directly implies that QFCs are at most exponentially more succinct than classical equivalent devices. Indeed, due to QFCs generality, this succinctness result transfers to other models of QFAs simulated by QFCs such as measure-only, measure-many, and reversible QFAs. Additionally, we here show how QFCs are also able to simulate Latvian and measure-only QFAs, thus providing size lower bounds even for these two models.

2. Preliminaries

2.1. Linear algebra

We quickly recall some notions of linear algebra, which are useful to describe quantum computing. For more details, we refer the reader to, e.g., [15,26]. The fields of real and complex numbers are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively.

Given a complex number $z = a + ib$, we denote its *real part*, *conjugate*, and *modulus* by $z_{\mathbb{R}} = a$, $z^* = a - ib$, and $|z| = \sqrt{zz^*}$, respectively. We denote by $\mathbb{C}^{n\times m}$ the set of $n \times m$ matrices with entries in $\mathbb{C}$. Given a matrix $M \in \mathbb{C}^{n\times m}$, for $1 \le i \le n$ and $1 \le j \le m$, we denote by M_{ij} its (i, j)th entry. The *transpose* of M is the matrix $M^T \in \mathbb{C}^{m\times n}$ satisfying $M^T{}_{ij} = M_{ji}$, while we let M^* be the matrix satisfying $M^*{}_{ij} = (M_{ij})^*$. The *adjoint* of M is the matrix $M^\dagger = (M^T)^*$.

For matrices $A, B \in \mathbb{C}^{n\times m}$, their *sum* is the $n \times m$ matrix $(A + B)_{ij} = A_{ij} + B_{ij}$. For matrices $C \in \mathbb{C}^{n\times m}$ and $D \in \mathbb{C}^{m\times r}$, their *product* is the $n \times r$ matrix $(CD)_{ij} = \sum_{k=1}^{m} C_{ik} D_{kj}$. For matrices $A \in \mathbb{C}^{n\times m}$ and $B \in \mathbb{C}^{p\times q}$, their *direct sum* and *Kronecker (or tensor) product* are the $(n + p) \times (m + q)$ and $np \times mq$ matrices defined, respectively, as follows:

$$A \oplus B = \begin{pmatrix} A & [\mathbf{0}] \\ [\mathbf{0}] & B \end{pmatrix}, \qquad A \otimes B = \begin{pmatrix} A_{11}B & \cdots & A_{1m}B \\ \vdots & \ddots & \vdots \\ A_{n1}B & \cdots & A_{nm}B \end{pmatrix},$$

where $[\mathbf{0}]$ denotes zero-matrices of suitable dimensions. When operations are allowed by matrix dimensions, we have $(A \otimes B) \cdot (C \otimes D) = AC \otimes BD$ and $(A \oplus B) \cdot (C \oplus D) = AC \oplus BD$.

A *Hilbert space* of dimension n is the linear space $\mathbb{C}^{1\times n}$ of n-dimensional complex row vectors equipped with sum and product by elements in $\mathbb{C}$, in which the *inner product* $\langle \varphi, \psi \rangle = \varphi \psi^\dagger$ is defined, for $\varphi, \psi \in \mathbb{C}^{1\times n}$. From now on, for the sake of simplicity, we will write $\mathbb{C}^n$ instead of $\mathbb{C}^{1\times n}$. The *norm* of a vector $\varphi \in \mathbb{C}^n$ is given by $\|\varphi\| = \sqrt{\langle \varphi, \varphi \rangle}$. For vectors $\varphi \in \mathbb{C}^n$ and $\psi \in \mathbb{C}^m$, their *direct sum* is the vector $\varphi \oplus \psi = (\varphi_1, \dots, \varphi_n, \psi_1, \dots, \psi_m) \in \mathbb{C}^{n+m}$.

We recall the following properties, for $\varphi, \psi, \xi, \zeta \in \mathbb{C}^n$ and $r \in \mathbb{R}$, which will turn out to be useful in our calculations:

$$\begin{array}{ll} \langle \varphi, \psi \rangle = \langle \psi, \varphi \rangle^* = \langle \psi^*, \varphi^* \rangle, & \langle \varphi + \psi, \xi \rangle = \langle \varphi, \xi \rangle + \langle \psi, \xi \rangle, \\ \langle r\varphi, \psi \rangle = r\langle \varphi, \psi \rangle = \langle \varphi, r\psi \rangle, & \langle \varphi \otimes \psi, \xi \otimes \zeta \rangle = \langle \varphi, \xi \rangle \langle \psi, \zeta \rangle, \\ \|\varphi - \psi\|^2 = \|\varphi\|^2 + \|\psi\|^2 - 2\langle \varphi, \psi \rangle_{\mathbb{R}}, & \langle \varphi \oplus \psi, \xi \oplus \zeta \rangle = \langle \varphi, \xi \rangle + \langle \psi, \zeta \rangle, \\ \left|\langle \varphi, \psi \rangle\right| \le \|\varphi\|\|\psi\| \quad \text{(Cauchy–Schwarz inequality)}, & \|\varphi \otimes \psi\| = \|\varphi\|\|\psi\|. \end{array}$$

The *angle between complex vectors φ and ψ* is defined as (see, e.g., [25]):

$$\mathrm{ang}(\varphi, \psi) = \arccos \frac{\langle \varphi, \psi \rangle_{\mathbb{R}}}{\|\varphi\|\|\psi\|}.$$

We say that φ is *orthogonal* to ψ if $\langle \varphi, \psi \rangle = 0$. Two subspaces $X, Y \subseteq \mathbb{C}^n$ are orthogonal if any vector in X is orthogonal to any vector in Y. In this case, the linear space generated by $X \cup Y$ is denoted by $X \dotplus Y$.