数理解析研究所講究録 1325 巻 2003 年 203-208

Kolmogorov complexity upper bound of probability in computable POVM measurement*

只木孝太郎 (Kohtaro Tadaki)

科学技術振興事業団 創造科学技術推進事業 今井量子計算機構プロジェクト
IMAI Quantum Computation and Information Project, ERATO,
Japan Science and Technology Corporation
E-mail: tadaki@qci.jst.go.jp

Abstract. We apply algorithmic information theory to quantum mechanics in order to shed light on an algorithmic structure which is inherent in quantum mechanics.

There are two equivalent ways to define the (classical) Kolmogorov complexity $K(s)$ of a given classical finite binary string s. In the standard way, $K(s)$ is defined as the length of the shortest input string for the universal self-delimiting Turing machine to output s. In the other way, we first introduce the so-called universal probability m, and then define $K(s)$ as $-\log_2 m(s)$ without using the concept of program-size. We generalize the universal probability to a matrix-valued function, and identify this function with a positive operator-valued measure (POVM), which describes the statistics of outcomes in a quantum measurement in the general setting. Based on this identification, we study a computable POVM measurement with countable measurement outcomes performed upon a finite dimensional quantum system. We show that, up to a multiplicative constant, $2^{-K(s)}$ is the upper bound for the probability of each measurement outcome s in such a quantum measurement. In what follows, the upper bound $2^{-K(s)}$ is shown to be optimal in a certain sense.

Keywords: algorithmic information theory, quantum measurement, universal probability, POVM, computability, quantum Kolmogorov complexity

1 Introduction

Algorithmic information theory is a theory of program-size complexity which has precisely the formal properties of classical information theory. In algorithmic information theory, the *program-size complexity* (or *Kolmogorov complexity*) $K(s)$ of a finite binary string s is defined as the length of the shortest binary input for the universal self-delimiting Turing machine to output s. The concept of program-size complexity plays an important role in characterizing the randomness of a finite or infinite binary string. In this paper we extend algorithmic information theory to quantum region in order to throw light upon an algorithmic feature of quantum mechanics.

1.1 Main result

We consider a quantum measurement performed upon a *finite dimensional* quantum system. A *positive operator-valued measure* (POVM) is a collection $\{E(m)\}$ of positive semi-definite Hermitian matrices which satisfies $\sum_m E(m) = I$ where I is the identity matrix.

*For the detail of this work see arXiv:quant-ph/0212071.

Each $E(m)$ is called a *POVM element* of this POVM. In general, the statistics of outcomes in a quantum measurement are described by a POVM $\{E(m)\}$. The label m refers to the measurement outcomes that may occur in the experiment. If the state of the quantum system is described by a normalized vector $|\psi\rangle$ immediately before the measurement, then the probability that result m occurs is given by $\langle\psi|E(m)|\psi\rangle$. On the other hand, if the ensemble of the states of the quantum system is described by a density matrix ρ immediately before the measurement, then the probability that result m occurs is given by $\mathrm{tr}(\rho E(m))$. A POVM measurement is a generalization of a familiar *projective measurement* which is described by an observable. The number of outcomes in a POVM measurement can be more than the dimension of the state space of the quantum system being measured, whereas the number of outcomes in a projective measurement cannot. In this paper, we relate an argument s of $K(s)$ to an outcome which may occur in the quantum measurement performed upon a finite dimensional quantum system. Since $K(s)$ is defined for all finite binary strings s, the countable outcomes have to be available in the corresponding quantum measurement. Thus we deal with a POVM measurement and not a projective measurement. (See e.g. [6, 7] for the details of POVM measurements.)

We say a POVM is *computable* if one can compute all its POVM elements to any desired degree of precision, and a POVM measurement is said to be *computable* if it is described by a computable POVM. Our main result is as follows: Let $\{R(s)\}$ be a computable POVM on a finite dimensional quantum system whose each element is labeled by a finite binary string. Then there exists an integer d such that, for all density matrix ρ and all finite binary string s,

$$K(s) - d \leq -\log_2 \mathrm{tr}(\rho R(s)), \qquad (1)$$

and also there exists a real number $c > 0$ such that, for all density matrix ρ and all finite binary string s,

$$\mathrm{tr}(\rho R(s)) \leq c\, P(s). \qquad (2)$$

Here $P(s)$ is the probability that the (classical) universal self-delimiting Turing machine halts and outputs s when it starts on the program tape filled with an infinite binary string generated by infinitely repeated tosses of a fair coin.

The inequality (1) states that, up to an additive constant, $K(s)$ is the lower bound for the $-\log_2$ of the probability of each measurement outcome s in a computable POVM measurement with countable outcomes performed upon a finite dimensional quantum system, i.e., $2^{-K(s)}$ is the upper bound for the probability of each outcome s up to a multiplicative constant. On the other hand, the inequality (2) states that, up to a multiplicative constant, $P(s)$ is the upper bound for the probability of each measurement outcome s in the same measurement. Note that the inequalities (1) and (2) are equivalent to each other.

The computability of a POVM measurement is thought to be intrinsic in the case where one performs the measurement in order to extract a valuable information from a quantum system because in such a case one has to be able to compute to any desired degree of precision all POVM elements of the POVM which describes the measurement. Hence, when one wants to extract a valuable information from a finite dimensional quantum system through a POVM measurement with countable outcomes, one faces with the limitation given by the inequality (1) (equivalently by (2)).

Especially, the inequality (2) is interesting. Since $P(s)$ is a probability which results from infinitely repeated tosses of a fair coin, $P(s)$ is just a classical probability. In the case where ρ is a pure state, the inequality (2) states that a purely quantum mechanical probability is bounded from above by a purely classical probability up to a multiplicative constant when one performs a computable POVM measurement with countable outcomes upon a finite dimensional quantum system in the pure state ρ.

The inequalitis (1) and (2) are obtained through a generalization of the so-called *universal probability* to a matrix-valued function. The Kolmogorov complexity $K(s)$ of a finite binary string s is originally defined using the concept of program-size. However, there is another way to define $K(s)$ without referring to such a con-

cept, that is, we first introduce a universal probability m, and then define $K(s)$ as $-\log_2 m(s)$. The universal probability is a function from the set of finite binary strings to the open interval $(0,1)$. In this paper we generalize the universal probability to a matrix-valued function while keeping the domain of definition the set of finite binary strings. Then this generalized universal probability is identified with an analogue of a POVM, and is called a *universal semi-POVM*. The inequalities (1) and (2) naturally follow from this identification.

1.2 Related works

Our aim is to generalize algorithmic information theory in order to understand the algorithmic feature of quantum mechanics. There are related works whose purpose is mainly to define the information content of an individual pure quantum state, i.e., to define the *quantum Kolmogorov complexity* of the quantum state [8, 1, 5], while we will not make such an attempt in this paper.

As we mentioned above, $K(s)$ can be defined as the $-\log_2$ of the universal probability without using the concept of program-size. [5] took this approach in order to define the information content of a pure quantum state. [5] first generalized the universal probability to a matrix-valued function $\boldsymbol{\mu}$, called *quantum universal semi-density matrix*. The $\boldsymbol{\mu}$ is a function which maps any positive integer N to an $N \times N$ positive semi-definite Hermitian matrix $\boldsymbol{\mu}(N)$ with its trace less than or equal to one. [5] proposed to regard $\boldsymbol{\mu}(N)$ as an analogue of a density matrix of a quantum system called *semi-density matrix*. Then, in order to measure the information content of a pure quantum state $|\psi\rangle \in \mathbb{C}^N$, [5] introduced the *quantum algorithmic entropies* $\underline{H}(|\psi\rangle)$ and $\overline{H}(|\psi\rangle)$ as $-\log_2\langle\psi|\boldsymbol{\mu}(N)|\psi\rangle$ and $-\langle\psi|(\log_2 \boldsymbol{\mu}(N))|\psi\rangle$, respectively. In general, the trace of a density matrix has to be equal to one. If the trace of $\boldsymbol{\mu}(N)$ is equal to one, then the quantity $\langle\psi|\boldsymbol{\mu}(N)|\psi\rangle$ in the definition of $\underline{H}(|\psi\rangle)$ has the meaning of the probability that the outcome is 'yes' when one performs the projective measurement described by the projector $|\psi\rangle\langle\psi|$ upon the quantum system in the mixed state $\boldsymbol{\mu}(N)$. However, the trace of $\boldsymbol{\mu}(N)$ is not equal to one for all but finitely many N because of its universality.

In quantum mechanics, what is represented by a matrix is either a quantum state or a measurement operator. In this paper we generalize the universal probability to a matrix-valued function in different way from [5], and identify it with an analogue of a POVM. We do not stick to defining the information content of a quantum state. Instead, we focus our thoughts on properly extending algorithmic information theory to quantum region while keeping an appealing feature of the theory. In this line we have the above inequalities (1) and (2).

In each of [8] and [1], the quantum Kolmogorov complexity of a qubit string was defined as a quantum generalization of the standard definition of classical Kolmogorov complexity; the length of the shortest input for the universal decoding algorithm U to output a finite binary string. Both [8] and [1] adopt the *universal quantum Turing machine* as a universal decoding algorithm U to output a quantum state in their definition. However, there is a difference between [8] and [1] with respect to the object which is allowed as an input to U. That is, [8] can only allow a classical binary string as an input, whereas [1] can allow any qubit string. The works [8], [1], and [5] are closely related to one another as shown in each of these works. In comparison with our work, since our work is, in essence, based on a generalization of the universal probability, the work [5] is more related to our work than the works [8] and [1]. These two works may be related to our work via the work [5].

2 Preliminaries

2.1 Notation

We start with some notation about numbers and matrices which will be used in this paper.

$\mathbf{N} \equiv \{0,1,2,3,\ldots\}$ is the set of natural numbers, and $\mathbf{N}^+$ is the set of positive integers. $\mathbf{Q}$ is the set of rational numbers, and $\mathbf{C}$ is the set of complex numbers. $\mathbf{C}_{\mathbf{Q}}$ is the set of the complex numbers in the form of $a+ib$ with $a, b \in \mathbf{Q}$. We define $-\log_2 0$ as ∞.

We fix N to be any one positive integer throughout this paper. $\mathbb{C}^N$ is the set of column vectors consist N complex numbers. For each $K \subset \mathbb{C}$, $M_N(K)$ is the set of the $N \times N$ matrices whose elements are in K. For each $A \in M_N(\mathbb{C})$, $A^\dagger$ is the *adjoint* of A, $\operatorname{tr} A$ is the *trace* of A, and $\|A\|$ is the *operator norm* of A. The *identity matrix* in $M_N(\mathbb{C})$ is denoted by I. $\mathrm{Her}(N)$ is the set of $N \times N$ Hermitian matrices. For each $A, B \in \mathrm{Her}(N)$, we write $A \leqslant B$ if $B - A$ is positive semi-definite, and write $A < B$ if $B - A$ is positive definite. Note that the relation $\leqslant$ on $\mathrm{Her}(N)$ is a partial order. We say ρ is a *density matrix* if $0 \leqslant \rho \in \mathrm{Her}(N)$ and $\mathrm{tr}(\rho) = 1$. $\mathrm{Her}_{\mathbb{Q}}(N)$ is the set of $N \times N$ Hermitian matrices whose elements are in $\mathbb{C}_{\mathbb{Q}}$.

Let S be any set, and let $f, g\colon S \to \mathrm{Her}(N)$. Then we write $f(x) = g(x) + O(1)$ if there is a real number $c > 0$ such that, for all $x \in S$, $\|f(x) - g(x)\| \leq c$. We also write $f(x) \sim g(x)$ if there is a real number $c > 0$ such that, for all $x \in S$, $c\,f(x) \leqslant g(x)$ and $c\,g(x) \leqslant f(x)$.

$\Sigma^* \equiv \{\lambda, 0, 1, 00, 01, 10, 11, 000, 001, 010, \ldots\}$ is the set of finite binary strings where λ denotes the *empty string*, and Σ^* is ordered as indicated. We identify any string in Σ^* with a natural number in this order, that is, we consider $\varphi\colon \Sigma^* \to \mathbf{N}$ such that $\varphi(s) = 1s - 1$ where the concatenation $1s$ of strings 1 and s is regarded as a dyadic integer, and then we identify s with $\varphi(s)$. For any $s \in \Sigma^*$, $|s|$ is the *length* of s. A subset S of Σ^* is called a *prefix-free set* if no string in S is a prefix of another string in S.

For each $F\colon \Sigma^* \to M_N(\mathbb{C})$, we say F is *computable* if there exists a total recursive function $G\colon \Sigma^* \times \mathbf{N} \to M_N(\mathbb{C}_{\mathbb{Q}})$ such that, for all $s \in \Sigma^*$ and all $k \in \mathbf{N}$, $\|F(s) - G(s,k)\| < 2^{-k}$.

2.2 Algorithmic information theory

In the following we review some definitions and results of algorithmic information theory [3, 4].

A *computer* is a partial recursive function $C\colon \Sigma^* \to \Sigma^*$ whose domain of definition is a prefix-free set. For each computer C and each $s \in \Sigma^*$, $K_C(s)$ is defined as $\min\{\,|p| \mid p \in \Sigma^* \,\&\, C(p) = s\,\}$. A computer U is said to be *optimal* if for each computer C there exists a constant $\mathrm{sim}(C)$ with the following property: if $C(p)$ is defined, then there is a p' for which $U(p') = C(p)$ and $|p'| \leq |p| + \mathrm{sim}(C)$. It is then shown that there exists a computer which is optimal. We choose any one optimal computer U as the standard one for use throughout the rest of this paper, and we define $K(s) \equiv K_U(s)$, which is referred to as the *information content* of s, the *program-size complexity* of s, or the *Kolmogorov complexity* of s. For each $s \in \Sigma^*$, $P(s)$ is defined by $P(s) \equiv \sum_{U(p)=s} 2^{-|p|}$. The class of computers is equal to the class of functions which are computed by *self-delimiting Turing machines.* A self-delimiting Turing machine has a program tape and a work tape. The machine starts with a finite binary string as input on its program tape. When the machine halts, the output string is put on the work tape. (For the details of self-delimiting Turing machine, see [3].) A self-delimiting Turing machine is called *universal* if it computes an optimal computer. Let M_U be a universal self-delimiting Turing machine which computes U. Then $P(s)$ is the probability that M_U halts and outputs s when M_U starts on the program tape filled with an infinite binary string generated by infinitely repeated tosses of a fair coin.

A universal probability is defined through the following two definitions.

Definition 2.1. *For any $r\colon \Sigma^* \to [0, \infty)$, we say that r is a lower-computable semi-measure if r satisfies the following two conditions:*

(i) $\sum_{s \in \Sigma^*} r(s) \leq 1$.

(ii) There exists a total recursive function $f\colon \mathbf{N} \times \Sigma^* \to \mathbf{Q}$ such that, for each $s \in \Sigma^*$, $\lim_{n\to\infty} f(n,s) = r(s)$ and $\forall n \in \mathbf{N}$ $f(n,s) \leq f(n+1,s)$.

Definition 2.2. *Let m be a lower-computable semi-measure. We say that m is a universal probability if for any lower-computable semi-measure r, there exists a real number $c > 0$ such that, for all $s \in \Sigma^*$, $c\,r(s) \leq m(s)$.*

Theorem 2.3. *Both $2^{-K(s)}$ and $P(s)$ are universal probabilities.*

By Theorem 2.3, we see that, for any universal probability m, $K(s) = -\log_2 m(s) + O(1)$. Thus it is possible to define $K(s)$ as $-\log_2 m(s)$

with any one universal probability m instead of $K_U(s)$.

3 Generalization of universal probability to POVM

In this section we generalize a universal probability to a matrix-valued function. Based on this generalization, we prove our main result: Theorem 3.7.

Definition 3.1. *We say R is a semi-POVM on Σ^* if R is a mapping from Σ^* to $\mathrm{Her}(N)$ which satisfies $0 \leqslant R(s)$ for all $s \in \Sigma^*$ and $\sum_{s\in\Sigma^*} R(s) \leqslant I$. We say R is a POVM on Σ^* if R is semi-POVM on Σ^* and $\sum_{s\in\Sigma^*} R(s) = I$.*

Let Q be a POVM on Σ^*. The POVM measurement described by Q is performed upon a finite dimensional quantum system, and gives one of countable measurement outcomes, which are represented by finite binary strings.

Given R: semi-POVM on Σ^*, it is easy to convert R into a POVM on Σ^* by appending an appropriate positive semi-definite matrix to R. Let $\Omega = \sum_{s\in\Sigma^*} R(s)$, and then we define $Q\colon \Sigma^* \to \mathrm{Her}(N)$ by $Q(\lambda) = I - \Omega$ and $Q(s') = R(s)$ for each $s \in \Sigma^*$ where s' is the successor of s. Then Q is a POVM on Σ^*. Thus a semi-POVM on Σ^* has a physical meaning in the same way as a POVM on Σ^*.

Definition 3.2. *We say R is a lower-computable semi-POVM if R is a semi-POVM on Σ^* and there exists a total recursive function $f\colon \mathbf{N} \times \Sigma^* \to \mathrm{Her}_{\mathbf{Q}}(N)$ such that, for each $s \in \Sigma^*$, $\lim_{n\to\infty} f(n,s) = R(s)$ and $\forall n \in \mathbf{N}\ f(n,s) \leqslant R(s)$.*

In the case where $N = 1$, Definition 3.2 exactly results in the definition of a lower-computable semi-measure.

Definition 3.3. *Let M be a lower-computable semi-POVM. We say that M is a universal semi-POVM if for each lower-computable semi-POVM R, there exists a real number $c > 0$ such that for all $s \in \Sigma^*$, $c\,R(s) \leqslant M(s)$.*

In the case where $N = 1$, Definition 3.3 exactly results in the definition of a universal probability. The use of the partial order $\leqslant$ for the purpose of generalizing lower-computable semi-measure and universal probability to matrix-valued functions is suggested in [5].

A universal semi-POVM may have a simple form as the following theorem says.

Theorem 3.4. *If m is a universal probability, then the mapping $\Sigma^* \ni s \longmapsto m(s)I$ is a universal semi-POVM.*

The following theorem is more general form of our main result.

Theorem 3.5. *Let m be a universal probability, and let R be a lower-computable semi-POVM. Then the following (i) and (ii) hold:*

(i) There exists $c > 0$ such that, for any normalized $|\psi\rangle \in \mathbb{C}^N$ and any $s \in \Sigma^*$,

$$\langle\psi|R(s)|\psi\rangle \leq c\,m(s).$$

(ii) There exists $c > 0$ such that, for any density matrix $\rho \in \mathrm{Her}(N)$ and any $s \in \Sigma^*$,

$$\mathrm{tr}(\rho R(s)) \leq c\,m(s).$$

Proof. It follows from Theorem 3.4 that (i) holds. Using (i) and the spectral decomposition of ρ, we have (ii). □

In order to make more clear the physical implication of Theorem 3.5, we restrict our attention to a POVM on Σ^* which is computable. Informally, a POVM on Σ^* is computable if and only if one can compute all its POVM elements to any desired degree of precision. Thus the computability of a POVM is thought to be inherent in the case where one wants to perform a well-controlled quantum measurement described by the POVM. Using the following lemma, we have our main result about a computable POVM.

Lemma 3.6. *Let R be a semi-POVM on Σ^*. If R is computable then R is a lower-computable semi-POVM.*

Theorem 3.7 (Main result). *Let R be a computable POVM on Σ^*. Then the following hold:*

(i) There exists $d \in \mathbb{N}$ such that, for any density matrix $\rho \in \mathrm{Her}(N)$ and any $s \in \Sigma^*$,

$$K(s) - d \leq -\log_2 \mathrm{tr}(\rho R(s)).$$

(ii) There exists $c > 0$ such that, for any density matrix $\rho \in \mathrm{Her}(N)$ and any $s \in \Sigma^*$,

$$\mathrm{tr}(\rho R(s)) \leq c\, P(s).$$

Proof. Theorem 3.7 immediately follows from Theorem 2.3, (ii) in Theorem 3.5, and Lemma 3.6. □

4 Optimality of universal semi-POVM

In this section we consider an optimality of a universal semi-POVM. By Theorem 2.3 and Theorem 3.4 we have the following theorem.

Theorem 4.1. *Let M be a universal semi-POVM. Then, for any density matrix $\rho \in \mathrm{Her}(N)$ and any $s \in \Sigma^*$,*

$$K(s) = -\log_2 \mathrm{tr}(\rho M(s)) + O(1),$$
$$P(s) \sim \mathrm{tr}(\rho M(s)).$$

Thus, if we can perform the POVM measurement described by a universal semi-POVM, then we can achieve the upper bound $P(s)$ (or $2^{-K(s)}$) in Theorem 3.7 up to a multiplicative constant. However any universal semi-POVM is not computable. Moreover we can show that there is no computable semi-POVM on Σ^* which can achieve the upper bound $P(s)$ (or $2^{-K(s)}$) up to a multiplicative constant. Instead, by the definition of universal semi-POVM, we have the following theorem. This theorem states that we can approximate any universal semi-POVM M by a recursive sequence $F_0, F_1, F_2, \ldots$ of POVMs which converges to the M from below as $n \to \infty$ in a certain sense.

Theorem 4.2. *For any universal semi-POVM M, there exists a sequence $F_0, F_1, F_2, \ldots$ such that*

(i) $F_n \colon \{s \mid s \leq n+1\} \to \mathrm{Her}_{\mathbb{Q}}(N)$ is a POVM for each $n \in \mathbb{N}$,

(ii) the mapping $\{(n,s) \mid s \leq n+1\} \ni (n,s) \longmapsto F_n(s)$ is a total recursive function, and

(iii) any given $\varepsilon > 0$, for all sufficiently large $n \in \mathbb{N}$, if $s \leq n$ and ρ is a density matrix then $0 \leq \mathrm{tr}(\rho M(s)) - \mathrm{tr}(\rho F_n(s)) < \varepsilon$.

Acknowledgments

The author is grateful to H. Imai and K. Matsumoto of the IMAI QCI project for their support.

References

[1] Berthiaume A., van Dam W., and Laplante S., Quantum Kolmogorov complexity, *J. Compute. System Sci.*, **63** (2001), 201–221.

[2] Calude C. S., *Information and Randomness: An Algorithmic Perspective*, 2nd Edition, Revised and Extended, Springer, Berlin, 2002.

[3] Chaitin G. J., A theory of program size formally identical to information theory, *J. Assoc. Comput. Mach.*, **22** (1975), pp.329–340.

[4] Chaitin G. J., Incompleteness theorems for random reals, *Adv. in Appl. Math.*, **8** (1987), pp.119–146.

[5] Gács P., Quantum algorithmic entropy, *J. Phys. A: Math. Gen.*, **34** (2001), pp.6859–6880.

[6] Nielsen M. A. and Chuang I. L., *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.

[7] Preskill J., *Quantum Computation*, 2000. Course notes available at URL: `http://www.theory.caltech.edu/people/preskill/ph229/` .

[8] Vitányi P. M. B., Quantum Kolmogorov complexity based on classical descriptions, *IEEE Trans. Inform. Theory*, **47** (2001), pp.2464–2479.