Journal of Uncertain Systems
Vol.5, No.x, pp.x-x, 2011
Online at: www.jus.org.uk

Expanding Algorithmic Randomness to the Algebraic Approach to Quantum Physics: Kolmogorov Complexity and Quantum Logics

Vladik Kreinovich*
Department of Computer Science, University of Texas at El Paso, El Paso, TX 79968, USA

Received 15 August 2009; Revised 23 February 2010

Abstract

Physicists usually assume that events with a very small probability cannot occur. Kolmogorov complexity formalizes this idea for non-quantum events. We show how this formalization can be extended to quantum events as well.

Keywords: Kolmogorov complexity, algorithmic randomness, quantum logic, algebraic approach to quantum physics

1 Algorithmic Randomness and Kolmogorov Complexity: Motivations and Definitions

Physicists' idea: events with a very small probability cannot occur. Physicists usually assume that events with a very small probability cannot occur. For example, in principle, due to the Brownian motion, it is possible that the kettle placed on a cold stove starts boiling. The probability of this event is positive but very small.

A mathematician would say that this event is possible but rare. However, a physicist would say that this event is simply not possible. It is desirable to formalize this intuition of physicists.

Kolmogorov's formalization. To formalize this intuition of physicists, Kolmogorov proposed a new definition of a random sequence, a definition that enables us to separate physically random binary sequences (like sequences that appear in coin flipping experiments or sequences that appear in quantum measurements) from sequence that follow some pattern.

Intuitively, if a sequence is random, it satisfies all the *probability laws*. A probability law is a statement S that is true with probability 1: $P(S) = 1$. So, to prove that a sequence is not random, we must show that it does not satisfy one of these laws.

Sequences that do not satisfy a given law S form a definable with probability 0. Vice versa, for every definable set C of probability 0, the statement of "not belonging" to this set is true with probability 1 and is, thus, a probability law.

Thus, we can say that a sequence s is not random if it belongs to some (definable) set of probability 0. Thus, in the 1960s, Kolmogorov and his student Martin-Löf defined a sequence s to be *random* if it does not belong to any definable set C of probability 0; see, e.g., [3] for details.

This definition is consistent. Indeed, every definable set C is defined by a finite sequence of symbols (its definition). Since there are countably many sequences of symbols, there are countably many definable sets C. So, the complement $-\mathcal{R}$ to the class of all $\mathcal{R}$ of all random sequences also has probability 0.

2 More Physically Adequate Versions of Kolmogorov Complexity and Their Applications

Towards a more physically adequate versions of Kolmogorov randomness. The original 1960s Kolmogorov's definition only explains why events with probability 0 do not happen. What we need is to formalize the physicists' intuition that events with very small probability cannot happen.

*Email: vladik@utep.edu (V. Kreinovich).

At first glance, there is a seemingly natural formalization of this intuition: there exists the "smallest possible probability" p_0 such that:

- if the computed probability p of some event is larger than p_0, then this event can occur, while
- if the computed probability p is $\le p_0$, the event cannot occur.

For example, we want to formalize the idea that a fair coin cannot fall heads 100 times in a row. Indeed, the probability 2^{-100} of this event is extremely small. So, if we select a threshold p_0 for which $p_0 \ge 2^{-100}$, we get the desired justification. However, on second glance, one can see that there is a problem with this approach. Indeed, every length-100 sequence of heads and tails has exactly the same probability. So, if we choose $p_0 \ge 2^{-100}$, we will thus exclude *all* sequences of 100 heads and tails. However, anyone can toss a coin 100 times and get *some* sequence. This proves that some such sequences are physically possible.

Comment. A similar argument is known as *Kyburg's lottery paradox*: In a big (e.g., state-wide) lottery, the probability of winning the Grand Prize is very small, so a reasonable person should not expect to win. However, some people do win big prizes.

Towards a new definition of randomness. A new more physically adequate definition of randomness was proposed in [4, 5, 6]. This definition is based on the following analysis of what it means to be normal.

Let us illustrate this analysis on the example of height. A person of height ≥ 6 ft is still normal. If instead of 6 ft, we consider 6 ft 1 in, 6 ft 2 in, etc., then at some point we will arrive at a height h_0 for which everyone taller than h_0 is abnormal. We are not sure what is this threshold value h_0, but we are sure that such a value h_0 exists.

This argument can be extended to a general situation. On the universal set U, we have sets $A_1 \supseteq A_2 \supseteq \ldots \supseteq A_n \supseteq \ldots$ for which $P(\cap A_n) = 0$. In the above example, A_1 is the set of all the people with height ≥ 6 ft, A_2 is the set of all the people with height ≥ 6 ft 1 in, etc. We know that there exists an N for which all elements of the set A_N are not random. Thus, we arrive at the following definition.

Definition 1. *A set $\mathcal{R} \subseteq U$ is called a* set of random elements *if for every definable sequence of sets A_n for which $A_n \supseteq A_{n+1}$ for all n and $P(\cap A_n) = 0$, there exists an integer N for which $A_N \cap \mathcal{R} = \emptyset$.*

Mathematical comment. For this definition to be precise, we need to define what "definable" means. Such a definition is straightforward. Let $\mathcal{L}$ be a theory, let $P(x)$ be a formula from the language of the theory $\mathcal{L}$, with one free variable x so that the set $\{x \,|\, P(x)\}$ is defined in $\mathcal{L}$. We will then call the set $\{x \,|\, P(x)\}$ *$\mathcal{L}$-definable*.

One potential problem of this definition is that it is a definition in a meta-language, and our objective is to be able to make mathematical statements about $\mathcal{L}$-definable sets. Thus, we must have a stronger theory $\mathcal{M}$ in which the class of all $\mathcal{L}$-definable sets is a countable set. One can prove that such $\mathcal{M}$ always exists; see, e.g., [4].

Coin example. Let us illustrate the above definition on the example of coin tossing. In this example, the universal set $U = \{\mathrm{H}, \mathrm{T}\}^{\mathbb{N}}$ is the set of all possible infinite sequences of Heads (H) and Tails (T). Here, A_n is the set of all the sequences that start with n heads. The sequence $\{A_n\}$ is decreasing and definable, and its intersection has probability 0. Therefore, for every set $\mathcal{R}$ of random elements of U, there exists an integer N for which $A_N \cap T = \emptyset$.

This means that if a sequence $s \in \mathcal{R}$ is random and starts with N heads, it must consist of heads only. In physical terms, it means that a random sequence cannot start with N heads. This is exactly what we wanted to formalize.

From random to typical (not abnormal). The above idea can be used beyond randomness, to formalize a more general idea that not all solutions to the physical equations are physically meaningful.

Indeed, when a cup breaks into pieces, the corresponding trajectories of molecules make physical sense. However, when we reverse all the velocities, we get pieces assembling themselves into a cup.

Newton's equations do not change when we simply reverse the time. Thus, from the mathematical viewpoint, this "reversed" solution satisfies the same physical equations as the original one. However, from the physical viewpoint, the reversed solution is clearly impossible.

A usual physicist's explanation is that the reversed process is non-physical since its initial conditions are "degenerate" ("abnormal"): once we modify the initial conditions even slightly, the pieces will no longer get together. How can we formalize this notion of abnormality?

Towards a new definition of non-abnormality. Let us go back to the same example that we used to formalize randomness. A person of height 6 ft is still normal. If instead of 6 ft, we consider 6 ft 1 in, 6 ft 2 in, etc., then exists a threshold value h_0 for which everyone taller than h_0 is abnormal. We are not sure what is this value h_0, but we are sure such such a value h_0 exists.

In general, on the universal set U, we have sets $A_1 \supseteq A_2 \supseteq \ldots \supseteq A_n \supseteq \ldots$ for which $\cap A_n = \emptyset$. In the above example, A_1 is the set of all the people with height ≥ 6 ft, A_2 is the set of all the people with height ≥ 6 ft 1 in, etc. Thus, we arrive at the following definition.

Definition 2. *A set $\mathcal{T} \subseteq U$ is called a* set of typical elements *if for every definable sequence of sets A_n for which $A_n \supseteq A_{n+1}$ for all n and $\cap A_n = \emptyset$, there exists an N for which $A_N \cap \mathcal{T} = \emptyset$.*

Coin example. Let us illustrate this definition on the coin example. In this example, we have the universal set $U = \{\mathrm{H}, \mathrm{T}\}^{\mathbb{N}}$, and A_n is the set of all the sequences that start with n heads and has a tail.

The sequence $\{A_n\}$ is decreasing and definable, and its intersection is empty. Therefore, for every set $\mathcal{T}$ of typical elements of U, there exists an integer N for which $A_N \cap T = \emptyset$. This means that if a sequence $s \in \mathcal{T}$ is "random" (i.e., has both heads and tails) and starts with N heads, it must consist of heads only.

In physical terms, this means that a random sequence cannot start with N heads. This is exactly what we wanted to formalize.

Consistency proof. Let us prove that this definition is consistent. Moreover, we will prove that for every $\varepsilon > 0$, there exists a set $\mathcal{T}$ of typical elements for which $\underline{P}(\mathcal{T}) \geq 1 - \varepsilon$.

Indeed, similarly to the case of definable sets, there are countably many definable sequences $\{A_n\}$: $\{A_n^{(1)}\}$, $\{A_n^{(2)}\}$, ... that satisfy the above conditions. For each such sequence k, we have $\cap A_n = \emptyset$ and thus, $P\left(A_n^{(k)}\right) \to 0$ as $n \to \infty$. Hence, there exists N_k for which $P\left(A_{N_k}^{(k)}\right) \leq \varepsilon \cdot 2^{-k}$. We can now take $\mathcal{T} \stackrel{\text{def}}{=} - \bigcup_{k=1}^{\infty} A_{N_k}^{(k)}$. For this set, since $P\left(A_{N_k}^{(k)}\right) \leq \varepsilon \cdot 2^{-k}$, we have

$$\overline{P}\left(\bigcup_{k=1}^{\infty} A_{N_k}^{(k)}\right) \leq \sum_{k=1}^{\infty} P\left(A_{N_k}^{(k)}\right) \leq \sum_{k=1}^{\infty} \varepsilon \cdot 2^{-k} = \varepsilon.$$

Hence, $\underline{P}(\mathcal{T}) = 1 - \overline{P}\left(\bigcup_{k=1}^{\infty} A_{N_k}^{(k)}\right) \geq 1 - \varepsilon$.

Ill-posed problems: in brief. Let us give an example that the above definition of typicalness is not just philosophically interesting, it is practically useful. As such an example, we will take ill-posed problems.

What are ill-posed problems and why are they useful? One of the main objectives of science is to provide *guaranteed* estimates for physical quantities and *guaranteed* predictions for these quantities. The problem is that both estimation and prediction are *ill-posed* in the sense that arbitrarily small (practically un-detectable) changes in measurement results can lead to drastic changes in the estimates and predictions.

For example, every measurement devices are inertial, hence they suppress high frequencies ω. So, for large ω, the signals $\varphi(x)$ and $\varphi(x) + \sin(\omega \cdot t)$ are indistinguishable.

At present, there are several approaches to solving this problem: statistical regularization (filtering), Tikhonov regularization (i.e., restricting ourselves to a certain class of signals, e.g., signals for which $|\dot{x}| \leq \Delta$), expert-based regularization, etc. The main problem with these approaches is that they provide no guarantee.

It turns out that if we restrict ourselves to "not abnormal" solutions, problems become well-posed. Let us explain this on the example of an ill-posed problem of estimating a state $s \in S$ based on the measurement results. In this context, a measurement can be formalized as a function f that maps every state $s \in S$ into an observation $r = f(s) \in R$. Usually, if we perform a sufficient number of measurements, then, theoretically, we

can reconstruct the state s as $s = f^{-1}(r)$. However, as we have mentioned, the problem is that small changes in r can lead to huge changes in s – i.e., that the inverse function f^{-1} is not continuous.

As the following result shows, the situation changes if we we restrict ourselves to "not abnormal" solutions.

Proposition 1. *Let S be a definably separable metric space, let $\mathcal{T}$ be a set of all not abnormal elements of S, and let $f : S \to R$ be a continuous 1-1 function. Then, the inverse mapping $f^{-1} : R \to S$ is continuous for every $r \in f(\mathcal{T})$.*

Proof. It is known that if a function f is continuous and 1-1 on a compact set, then the inverse function f^{-1} is also continuous. To, to prove our result, it is sufficient to show that the set $\mathcal{T}$ can be included in some compact set.

Let us recall that a set X is compact if and only if it is closed and for every ε, it has a finite ε-net, i.e., a finite set $\{s_1, \ldots, s_N\}$ such that every point $x \in X$ is ε-close to one of the values s_i. It is thus sufficient to prove that for every for every ε, the set $\mathcal{T}$ has a finite ε-net. Then, its closure $\overline{\mathcal{T}}$ will be the desired compact set.

Indeed, the space S is definably separable, meaning that there exists a definable sequence $s_1, \ldots, s_n, \ldots$ which is everywhere dense in S. Let us now take $A_n \stackrel{\text{def}}{=} -\bigcup_{i=1}^{n} B_\varepsilon(s_i)$, where $B_\varepsilon(s) \stackrel{\text{def}}{=} \{x : d(x,s) \le \varepsilon\}$ denotes an ball of radius ε with a center in a point s.

Since s_i are everywhere dense, we have $\cap A_n = \emptyset$. Hence, there exists N for which $A_N \cap \mathcal{T} = \emptyset$. Since $A_N = -\bigcup_{i=1}^{N} B_\varepsilon(s_i)$, this means that $\mathcal{T} \subseteq \bigcup_{i=1}^{N} B_\varepsilon(s_i)$. Hence, the set $\{s_1, \ldots, s_N\}$ is the desired ε-net for $\mathcal{T}$. The proposition is proven.

Another practical use of algorithmic randomness: when to stop an iterative algorithm. In numerical mathematics, we often know an iterative process whose results x_k are known to converge to the desired solution x, but we do not know when to stop to guarantee that $d_X(x_k, x) \le \varepsilon$.

A usual heuristic approach is to stop when $d_X(x_k, x_{k+1}) \le \delta$ for some $\delta > 0$. For example, in physics, we can use linear, quadratic, etc. approximations. Usually, we assume that if the 2nd order terms are small, then we can use the linear (1st order) expression as an approximation.

The use of typicalness enables us to provide a guaranteed stopping criterion. Indeed, let $\{x_k\} \in S$, k be an integer, and $\varepsilon > 0$ a real number. We say that the value x_k is *ε-accurate* if $d_X(x_k, \lim x_p) \le \varepsilon$.

Let $d \ge 1$ be an integer. By a *stopping criterion*, we mean a function $c : X^d \to \mathbb{R}_0^+$ that satisfies the following two properties:

- if $\{x_k\} \in S$, then $c(x_k, \ldots, x_{k+d-1}) \to 0$; and
- if for some $\{x_n\} \in S$ and k, $c(x_k, \ldots, x_{k+d-1}) = 0$, then $x_k = \ldots = x_{k+d-1} = \lim x_p$.

Proposition 2. [4] *Let c be a stopping criterion. Then, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that if $c(x_k, \ldots, x_{k+d-1}) \le \delta$, and the sequence $\{x_n\}$ is not abnormal, then x_k is ε-accurate.*

3 Extension to Quantum Physics: Motivations, Definitions, and Results

Need to extend algorithmic randomness to quantum physics. The above definitions of algorithmic randomness and typicalness assume that we have a set (of possible states) and a probability measure on the set of all the states. In other words, these definitions cover only classical (non-quantum) physics.

In quantum physics, for each measurable quantity, we also have a probability distribution, but in general, there is no single probability distribution describing a given quantum state. Instead, for each binary (yes-no) observable a, we have the probability $m(a)$ of the "yes" answer.

Traditionally, logic is understood as the study of all yes-no statements, i.e., statements that can be either true or false. Correspondingly, the set of all binary observables in quantum physics is usually called a *quantum logic* L.

In the traditional formulation of quantum mechanics (see, e.g., [1, 2]), states are described by complex-valued functions $\psi(x)$ (called *wave functions*). The probability of finding a particle of for which $\int |\psi(x)|^2\,dx < \infty$. In mathematical terms, the set of all such functions forms a a *Hilbert space* H. In these terms, observables correspond to (closed) subspaces of this Hilbert space: every observable can be identified with the set of all the wave functions for which this observable returns "true" with probability 1. (However, it turned out be useful to consider more general (non-Hilbert) quantum logics as well.)

It is therefore natural to provide the following modification of the above definition.

Toward a natural extension of randomness to quantum logics. Let us recall in the non-quantum case, a set $\mathcal{T} \subseteq U$ is called a set of typical elements if for every definable sequence of sets A_n for which $A_n \supseteq A_{n+1}$ for all n and $\cap A_n = \emptyset$, there exists an N for which $A_N \cap \mathcal{T} = \emptyset$.

In this definition, a set A is *possible* if $A \cap \mathcal{T} \neq \emptyset$ and *impossible* if $A \cap \mathcal{T} = \emptyset$.

In quantum logic:

- a natural analogue of the universal set U is the quantum logic, the set L,
- a natural analogue of the subset relation $\supseteq$ is the ordering relation $\geq$, subset relation between the corresponding subspaces,
- an analogue of the intersection is the meet $\wedge$ – intersection of the subspaces, and
- a natural analogue of the empty set is the 0 element (= an observable that is always false).

This analogy enables us to extend the above definition to quantum logics that do not necessarily come from a Hilbert space.

Definition 3. *An element $T \in L$ is called* largest-typical *if for every definable sequence $A_n \in L$ for which $A_n \geq A_{n+1}$ for all n and $\wedge A_n = 0$, there exists an integer N for which $A_N \wedge T = 0$.*

Comment. Similarly to the non-quantum case, we say that A is *possible* if $A \wedge T \neq 0$ and *impossible* if $A \wedge T = 0$.

Towards a consistency result. Is this definition consistent? We would like to prove that for every $\varepsilon > 0$, there exists a largest-typical element T for which $m(T) \geq 1 - \varepsilon$.

To prove this result, we assume that L is a complete ortholattice such that:

- if $A_n \geq A_{n+1}$, then $A_n \to \wedge A_n$;
- lattice operations $\vee$ and $\wedge$ are continuous;
- the function $m : L \to [0, 1]$ is continuous.

Comment. It should be mentioned that these assumptions are non-trivial: e.g., for subspaces of the 2-dimensional space $\mathbb{R}^2$, the join operation $\vee$ is not continuous. Indeed, by definition, the join $a \vee b$ is the smallest element that is larger than both a and b. For linear subspaces a and b, this means that is the smallest linear space that contains both a and b – i.e., the linear span of the union

So, if a is a straight line, and b_n is a line at an angle $\alpha_n = \dfrac{1}{n} \to 0$ from a, then $a \vee b_n = \mathbb{R}^2$ for all n, so $a \vee b_n \to \mathbb{R}^2$. However, in the limit, $b_n \to a$ and thus, $a \vee b_n = \mathbb{R}^2 \not\to a \vee a = a$.

Consistency proof. The main idea behind this proof is the same as before:

- there exists countably many definable sequences $\{A_n\}$:

$$\{A_n^{(1)}\}, \quad \{A_n^{(2)}\}, \quad \ldots;$$

- so, we can take $T \stackrel{\text{def}}{=} -\bigvee_{k=1}^{\infty} A_{N_k}^{(k)}$ for some N_k.

However, in the quantum case, we cannot directly apply the original proof. Indeed, the original proof used the fact that $P(A \vee B) \le P(A) + P(B)$, but in quantum logic, we may have $m(A \vee B) > m(A) + m(B)$.

So, our new idea is to select N_k for which

$$m\left(A_{N_1}^{(1)} \vee \ldots \vee A_{N_k}^{(k)}\right) < \varepsilon.$$

Let us show how this can be done.

Let us assume that we have already selected the values $N_1, \ldots, N_k$ for which

$$m\left(A_{N_1}^{(1)} \vee \ldots \vee A_{N_k}^{(k)}\right) < \varepsilon.$$

Since $A_n^{(k+1)} \to 0$ and the join operation $\vee$ is continuous, we have

$$A_{N_1}^{(1)} \vee \ldots \vee A_{N_k}^{(k)} \vee A_n^{(k+1)} \to A_{N_1}^{(1)} \vee \ldots \vee A_{N_k}^{(k)}.$$

Since m is continuous, we have

$$m\left(A_{N_1}^{(1)} \vee \ldots \vee A_{N_k}^{(k)} \vee A_n^{(k+1)}\right) \to m\left(A_{N_1}^{(1)} \vee \ldots \vee A_{N_k}^{(k)}\right) < \varepsilon.$$

So, there exists a value N_{k+1} for which

$$m\left(A_{N_1}^{(1)} \vee \ldots \vee A_{N_k}^{(k)} \vee A_{N_{k+1}}^{(k+1)}\right) < \varepsilon.$$

In the limit, we have $m(-T) = m\left(\bigvee\limits_{k=1}^{\infty} A_{N_k}^{(k)}\right) \le \varepsilon$, hence $m(T) \ge 1 - \varepsilon$. The statement is proven.

Acknowledgments

This work was supported in part by the National Science Foundation grants HRD-0734825 and DUE-0926721, and by Grant 1 T36 GM078000-01 from the National Institutes of Health.

References

[1] Auletta, G., *Foundations and Interpretation of Quantum Mechanics*, World Scientific, Singapore, 2000.

[2] Feynman, R., R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.

[3] Li, M., and P. M. B. Vitanyi, *An Introduction to Kolmogorov Complexity*, Springer, 2009.

[4] Kreinovich, V., and A. M. Finkelstein, Towards applying computational complexity to foundations of physics, *Notes of Mathematical Seminars of St. Petersburg Department of Steklov Institute of Mathematics*, vol.316, pp.63–110, 2004; also in: *Journal of Mathematical Sciences*, vol.134, no.5, pp.2358–2382, 2006.

[5] Kreinovich, V., Towards a more physically adequate definition of randomness: a topological approach, *Journal of Uncertain Systems*, vol.1, no.4, pp.256–266, 2007.

[6] Kreinovich, V., Toward formalizing non-monotonic reasoning in physics: the use of Kolmogorov complexity", *Revista Iberoamericana de Inteligencia Artificial*, vol.41, pp.4–20, 2009.

[7] Kreinovich, V., Expanding algorithmic randomness to the algebraic approach to quantum physics: Kolmogorov complexity and quantum logics, *Abstracts of the BLAST conference (Boolean Algebras, Algebraic Logic, Quantum Logic, Universal Algebra, Set Theory, Set-theoretic Topology and Point-free Topology*, Las Cruces, New Mexico, August 10–14, 2009, pp.27–29.