Journal of Uncertain Systems
Vol.8, No.3, pp.227-232, 2014
Online at: www.jus.org.uk

If Many Physicists are Right and No Physical Theory is Perfect, Then the Use of Physical Observations can Enhance Computations

Michael Zakharevich[1], Olga Kosheleva[2,*]
[1] *Align Technology, Inc., 2560 Orchard Parkway, San Jose, California 95131*
[2] *University of Texas at El Paso, El Paso, Texas 79968, USA*

Received 27 June 2013; Revised 10 August 2013

Abstract

The questions of what is computable in the physical world are usually analyzed in the context of a physical theory – e.g., what is computable in Newtonian physics, what is computable in quantum physics, etc. Many physicists believe that no physical theory is perfect, i.e., that no matter how many observations support a physical theory, inevitably, new observations will come which will require this theory to be updated. We show, somewhat unexpectedly, that if such a no-perfect-theory principle is true, then the use of physical data can drastically enhance computations.

Keywords: computability, physical theories, Kolmogorov complexity

1 Formulation of the Problem

What is computable? – an important question. Since the pioneering 1930s works of Turing, researchers are interested in finding out what is computable. In the 1930s (and even in the 1950s, when computer were still in their infancy), this question was more philosophical than practical. Nowadays, with the current exponential increase in computation speed, the difference between computable "in principle" and computable right now is less drastic, it may be a difference of a few years – not surprisingly, scientists are seriously talking about the possibility of computers exceeding human ability to process information. Because of this practicality, the question of what is computable becomes more and more of practical interest.

At present, this question is usually analyzed within a physical theory. At present, the question of what is computable and what is not computable is analyzed within a given physical theory: e.g., within Newtonian physics, within quantum physics, etc.; see, e.g., [8].

No physical theory is perfect: a widely spread physicists' belief. If we find an answer to the computability question within a given physical theory, is this answer fully satisfactory? So far, in the history of physics, no matter how good a physical theory, no matter how good its accordance with observations, eventually, new observations appear which are not fully consistent with the original theory – and thus, a theory needs to be modified. For example, for several centuries, Newtonian physics seems to explain all observable facts – until later, quantum (and then relativistic) effects were discovered which required changes in physical theories.

Because of this history, many physicists believe that every physical theory is approximate – no matter how sophisticated a theory, no matter how accurate its current predictions, inevitably new observations will surface which would require a modification of this theory; see, e.g., [1].

*Corresponding author.
Emails: ymzakharevich@yahoo.com (M. Zakharevich), olgak@utep.edu (O. Kosheleva).

How does this belief affect computations? At first glance, the fact that no theory is perfect makes the question of what is computable rather hopeless: no matter how seriously we analyze computability within a given physical theory, eventually, this theory will turn out to be, strictly speaking, false – and thus, our analysis of what is computable will have to be redone.

In this paper, we show, however, that in spite of this seeming hopelessness, some important answers to the question of what is computable can be deduced simply from the fact no physical theory is perfect – namely, in this case, computations can be enhanced in comparison with the usual Turing machine computability.

Comment. Some preliminary results of this paper appeared in [3].

2 How to Describe, in Precise Terms, that No Physical Theory is Perfect

Discussion. The statement that no physical theory is perfect means that no matter what physical theory we have, eventually there will be observations which violate this theory. To formalize this statement, we need to formalize what are observations and what is a theory.

What are observation? Each observation can be represented, in the computer, as a sequence of 0s and 1s; actually, in many cases, the sensors already produce the signal in the computer-readable form, as a sequence of 0s and 1s. From this viewpoint, all past and future observations form a (potentially) infinite sequence $\omega = \omega_1\omega_2\ldots$ of 0s and 1s, $\omega_i \in \{0,1\}$.

What is a physical theory from the viewpoint of our problem: a set of sequences. A physical theory may be very complex, but all we care about is which sequences of observations ω are consistent with this theory and which are not. In other words, for our purposes, we can identify a physical theory T with the set of all sequences ω which are consistent with this theory.

Not every set of sequences corresponds to a physical theory: the set T must be non-empty and definable. Not every set of sequences comes from a physical theory. First, a physical theory must have at least one possible sequence of observations, i.e., the set T must be *non-empty*.

Second, a theory – and thus, the corresponding set – must be described by a finite sequence of symbols in an appropriate language. Sets which are uniquely by (finite) formulas are known as *definable*. Thus, the set T must be definable.

Since at any moment of time, we only have finitely many observations, the set T must be closed. Another property of a physical theory comes from the fact that at any given moment of time, we only have finitely many observations, i.e., we only observe finitely many bits. From this viewpoint, we say that observations $\omega_1,\ldots,\omega_n$ are consistent with the theory T if there is a continuing infinite sequence which is consistent with this theory, i.e., which belongs to the set T.

The only way to check whether an infinite sequence $\omega = \omega_1\omega_2\ldots$ is consistent with the theory is to check that for every n, the sequences $\omega_1\ldots\omega_n$ are consistent with the theory T. In other words, we require that for some every infinite $\omega = \omega_1\omega_2\ldots$,

- if for every n, the sequence $\omega_1\ldots\omega_n$ is consistent with the theory T, i.e., if for every n, there exists a sequence $\omega^{(n)} \in T$ which has the same first n bits as ω, i.e., for which $\omega_i^{(n)} = \omega_i$ for all $i = 1,\ldots,n$,
- then the sequence ω itself should be consistent with the theory, i.e., this infinite sequence should also belong to the set T.

From the mathematical viewpoint, we can say that the sequences $\omega^{(n)}$ converge to ω: $\omega^{(n)} \to \omega$ (or, equivalently, $\lim \omega^{(n)} = \omega$), where convergence is understood in terms of the usual metric on the set of all infinite sequences $d(\omega,\omega') \stackrel{\text{def}}{=} 2^{-N(\omega,\omega')}$, where $N(\omega,\omega') \stackrel{\text{def}}{=} \max\{k : \omega_1\ldots\omega_k = \omega'_1\ldots\omega'_k\}$.

In general, if $\omega^{(m)} \to \omega$ in the sense of this metric, this means that for every n, there exists an integer ℓ such that for every $m \geq \ell$, we have $\omega_1^{(m)}\ldots\omega_n^{(m)} = \omega_1\ldots\omega_n$. Thus, if $\omega^{(m)} \in T$ for all m, this means that for every n, a finite sequence $\omega_1 \dots \omega_n$ can be a part of an infinite sequence which is consistent with the theory T. In view of the above, this means that $\omega \in T$.

In other words, if $\omega^{(m)} \to \omega$ and $\omega^{(m)} \in T$ for all m, then $\omega \in T$. So, the set T must contain all the limits of all its sequences. In topological terms, this means that the set T must be *closed*.

A physical theory must be different from a fact and hence, the set T must be nowhere dense. The assumption that we are trying to formalize is that no matter how many observations we have which confirm a theory, there eventually will be a new observation which is inconsistent with this theory. In other words, for every finite sequence $\omega_1 \dots \omega_n$ which is consistent with the set T, there exists a continuation of this sequence which does not belong to T. The opposite would be if all the sequences which start with $\omega_1 \dots \omega_n$ belong to T; in this case, the set T will be *dense* in this set. Thus, in mathematical terms, the statement that every finite sequence which is consistent with T has a continuation which is not consistent with T means that the set T is *nowhere dense*.

Resulting definition of a theory. By combining the above properties of a set T which describes a physical theory, we arrive at the following definition.

Definition 1. *By a* physical theory*, we mean a non-empty closed nowhere dense definable set T.*

Mathematical comment. To properly define what is definable, we need to have a consistent formal definition of definability. In this paper, we follow a natural definition from [4, 5] – which is reproduced in the Appendix.

Formalization of the principle that no physical theory is perfect. In terms of the above notations, the no-perfect-theory principle simply means that the infinite sequence ω (describing the results of actual observations) is not consistent with any physical theory, i.e., that the sequence ω does not belong to any physical theory T. Thus, we arrive at the following definition.

Definition 2. *We say that an infinite binary sequence ω is* consistent with the no-perfect-theory principle *if the sequence ω does not belong to any physical theory (in the sense of Definition 1).*

Comment. Are there such sequences in the first place? Our answer is yes. Indeed, by definition, we want a sequence which does not belong to a union of all definable physical theories. Every physical theory is closed nowhere dense set. Every definable set is defined by a finite sequence of symbols, so there are no more than countably many definable theories. Thus, the union of all definable physical theories is contained in a union of countably many closed nowhere dense sets. Such sets are knows as *meager* (or *Baire first category*); it is known that the set of all infinite binary sequences is not meager. Thus, there are sequences who do not belong to the above union – i.e., sequences which are consistent with the no-perfect-theory principle; see, e.g., [2, 7].

3 How to Describe When Access to Physical Observations Enhances Computations

How to describe general computations. Each computation is a solution to a well-defined problem. As a result, each bit in the resulting answer satisfies a well-defined mathematical property. All mathematical properties can be described, e.g., in terms of Zermelo-Fraenkel theory ZF, the standard formalization of set theory. For each resulting bit, we can formulate a property P which is true if and only if this bit is equal to 1. In this sense, each bit in each computation result can be viewed as the truth value of some statement formulated in ZF. Thus, our general ability to compute can be described as the ability to (at least partially) compute the sequence of truth values of all statements from ZF.

All well-defined statements from ZF can be numbered, e.g., in lexicographic order. Let α_n denote the truth value of the n-th ZF statement, and let $\alpha = \alpha_1 \dots \alpha_n \dots$ denote the infinite sequence formed by these truth values. In terms of this sequence, our ability to compute is our ability to compute the sequence α.

Kolmogorov complexity as a way to describe what is easier to compute. We want to analyze whether the use of physical observations (i.e., of the sequence ω analyzed in the previous section) can simplify computations. A natural measure of easiness-to-compute was invented by Kolmogorov, the founder of modern probability theory, when he realized that in the traditional probability theory, there is no formal way to distinguish between:

- finite sequences which come from observing from truly random processes, and
- orderly sequences like $0101\ldots01$.

Kolmogorov noticed that an orderly sequence $0101\ldots01$ can be computed by a short program, while the only way to compute a truly random sequence $0101\ldots$ is to have a print statement that prints this sequence. He suggested to describe this differences by introducing what is now known as *Kolmogorov complexity* $K(x)$ of a finite sequence x: the shortest length of a program (in some programming language) which computes the sequence x.

- For an orderly sequence x, the Kolmogorov complexity $K(x)$ is much smaller than the length $\mathrm{len}(x)$ of this sequence: $K(x) \ll \mathrm{len}(x)$.
- For a truly random sequence x, we have $K(x) \approx \mathrm{len}(x)$; see, e.g., [6].

The smaller the difference $\mathrm{len}(x) - K(x)$, the more we are sure that the sequence x is truly random.

Relative Kolmogorov complexity as a way to describe when using an auxiliary sequence simplifies computations. The usual notion of Kolmogorov complexity provides the complexity of computing x "from scratch". A similar notion of the *relative Kolmogorov complexity* $K(x\,|\,y)$ can be used to describe the complexity of computing x when a (potentially infinite) sequence y is given. This relative complexity is based on programs which are allowed to use y as a subroutine, i.e., programs which, after generating an integer n, can get the n-th bit y_n of the sequence y by simply calling y. When we compute the length of such programs, we just count the length of the call, not the length of the auxiliary program which computes y_n – just like when we count the length of a C++ program, we do not count how many steps it takes to compute, e.g., $\sin(x)$, we just count the number of symbols in this function call. The relative Kolmogorov complexity is then defined as the shortest length of such a y-using program which computes x.

Clearly, if x and y are unrelated, having access to y does not help in computing x, so $K(x\,|\,y) \approx K(x)$. On the other hand, if x coincides with y, then the relative complexity $K(x\,|\,y)$ is very small: all we need is a simple for-loop, in which we call for each bit y_i, $i = 1, \ldots, n$, and print this bit right away.

Resulting reformulation of our question. In terms of relative Kolmogorov complexity, the question of whether observations enhance computations is translated into checking whether $K(\alpha_1\ldots\alpha_n\,|\,\omega) \approx K(\alpha_1\ldots\alpha_n)$ (in which case there is no enhancement) or whether $K(\alpha_1\ldots\alpha_n\,|\,\omega) \ll K(\alpha_1\ldots\alpha_n)$ (in which case there is a strong enhancement). The larger the difference $K(\alpha_1\ldots\alpha_n) - K(\alpha_1\ldots\alpha_n\,|\,\omega)$, the larger the enhancement.

4 Main Result: Enhancement is Possible

Let us show that under the no-perfect-theory principle, observations do indeed enhance computations.

Proposition. *Let α be a sequence of truth values of ZF statements, and let ω be an infinite binary sequence which is consistent with the no-perfect-theory principle. Then, for every integer $C > 0$, there exists an integer n for which $K(\alpha_1\ldots\alpha_n\,|\,\omega) < K(\alpha_1\ldots\alpha_n) - C$.*

Comment. In other words, in principle, we can have an arbitrary large enhancement.

Proof. Let us fix an integer C. To prove the desired property for this C, let us prove that the set T of all the sequences which do not satisfy this property, i.e., for which $K(\alpha_1 \ldots \alpha_n \,|\, \omega) \geq K(\alpha_1 \ldots \alpha_n) - C$ for all n, is a physical theory in the sense of Definition 1. For this, we need to prove that this set T is non-empty, closed, nowhere dense, and definable. Then, from Definition 2, it will follow that the sequence ω does not belong to this set and thus, that the conclusion of Proposition 1 is true.

The set T is clearly non-empty: it contains, e.g., a sequence $\omega = 00\ldots 0\ldots$ which does not affect computations. The set T is also clearly definable: we have just defined it.

Let us prove that the set T is closed. For that, let us assume that $\omega^{(m)} \to \omega$ and $\omega^{(m)} \in T$ for all m. We then need to prove that $\omega \in T$. Indeed, let us fix n, and let us prove that $K(\alpha_1 \ldots \alpha_n \,|\, \omega) \geq K(\alpha_1 \ldots \alpha_n) - C$. We will prove this by contradiction. Let us assume that $K(\alpha_1 \ldots \alpha_n \,|\, \omega) < K(\alpha_1 \ldots \alpha_n) - C$. This means that there exists a program p of length $\mathrm{len}(p) < K(\alpha_1 \ldots \alpha_n) - C$ which uses ω to compute $\alpha_1 \ldots \alpha_n$. This program uses only finitely many bits of ω; let B be the largest index of these bits. Due to $\omega^{(m)} \to \omega$, there exists M for which, for all $m \geq M$, the first B bits of $\omega^{(m)}$ coincide with the first B bits of the sequence ω. Thus, the same program p will work exactly the same way – and generate the sequence $\alpha_1 \ldots \alpha_n$ – if we use $\omega^{(m)}$ instead of ω. But since $\mathrm{len}(p) < K(\alpha_1 \ldots \alpha_n) - C$, this would means that the shortest length $K(\alpha_1 \ldots \alpha_n \,|\, \omega^{(m)})$ of all the programs which use $\omega^{(m)}$ to compute $\alpha_1 \ldots \alpha_n$ also satisfies the inequality $K(\alpha_1 \ldots \alpha_n \,|\, \omega^{(m)}) < K(\alpha_1 \ldots \alpha_n) - C$. This inequality contradicts to our assumption that $\omega^{(m)} \in T$ and thus, that $K(\alpha_1 \ldots \alpha_n \,|\, \omega^{(m)}) \geq K(\alpha_1 \ldots \alpha_n) - C$. The contradiction proves that the set T is indeed closed.

Let us now prove that the set T is nowhere dense, i.e., that for every finite sequence $\omega_1 \ldots \omega_m$, there exists a continuation ω which does not belong to the set T. Indeed, as such a continuation, we can simply take a sequence $\omega = \omega_1 \ldots \omega_m \alpha_1 \alpha_2 \ldots$ obtained by appending α at the end. For this new sequence, computing $\alpha_1 \ldots \alpha_n$ is straightforward: we just copy the values α_i from the corresponding places of the new sequence ω. Here, the relative Kolmogorov complexity $K(\alpha_1 \ldots \alpha_n \,|\, \omega)$ is very small and is, thus, much smaller than the complexity $K(\alpha_1 \ldots \alpha_n)$ which – since ZF is not decidable – grows fast with n.

The proposition is proved.

Acknowledgments

The authors are thankful to Vladik Kreinovich, Luc Longpré, and Sergei Soloviev for valuable discussions.

References

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

[2] Jalal-Kamali, A., Nebesky, O., Durcholz, M.H., Kreinovich, V., and L. Longpré, Towards a "generic" notion of genericity: from 'typical' and 'random' to meager, shy, etc., *Journal of Uncertain Systems*, vol.6, no.2, pp.104–113, 2012.

[3] Kosheleva, O.M., and S.V. Soloviev, On the logic of using observable events in decision making, *Proceedings of the IX National USSR Symposium on Cybernetics*, pp.49–51, 1981.

[4] 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.

[5] Kreinovich, V., and A.M. Finkelstein, Towards applying computational complexity to foundations of physics, *Journal of Mathematical Sciences*, vol.134, no.5, pp.2358–2382, 2006.

[6] Li, M., and P.M.B. Vitányi, *An Introduction to Kolmogorov Complexity*, Springer-Verlag, Berlin, Hiedelberg, New York, 2008.

[7] Oxtoby, J.C., *Measure and Category: A Survey of the Analogies between Topological and Measure Spaces*, Springer Verlag, New York, Heidelberg, Berlin, 1980.

[8] Papadimitriou, C., *Computational Complexity*, Addison Welsey, Reading, Massachusetts, 1994.

A Appendix: A Formal Definition of Definable Sets

Definition A1. *Let $\mathcal{L}$ be a theory, and let $P(x)$ be a formula from the language of the theory $\mathcal{L}$, with one free variable x for which the set $\{x \,|\, P(x)\}$ is defined in the theory $\mathcal{L}$. We will then call the set $\{x \,|\, P(x)\}$* $\mathcal{L}$-definable.

Crudely speaking, a set is $\mathcal{L}$-definable if we can explicitly *define* it in $\mathcal{L}$. The set of all real numbers, the set of all solutions of a well-defined equation, every set that we can describe in mathematical terms: all these sets are $\mathcal{L}$-definable.

This does not mean, however, that *every* set is $\mathcal{L}$-definable: indeed, every $\mathcal{L}$-definable set is uniquely determined by formula $P(x)$, i.e., by a text in the language of set theory. There are only denumerably many words and therefore, there are only denumerably many $\mathcal{L}$-definable sets. Since, e.g., in a standard model of set theory ZF, there are more than denumerably many sets of integers, some of them are thus not $\mathcal{L}$-definable.

Our objective is to be able to make mathematical statements about $\mathcal{L}$-definable sets. Therefore, in addition to the theory $\mathcal{L}$, we must have a stronger theory $\mathcal{M}$ in which the class of all $\mathcal{L}$-definable sets is a set – and it is a countable set.

Denotation. *For every formula F from the theory $\mathcal{L}$, we denote its Gödel number by $\lfloor F \rfloor$.*

Comment. A Gödel number of a formula is an integer that uniquely determines this formula. For example, we can define a Gödel number by describing what this formula will look like in a computer. Specifically, we write this formula in LaTeX, interpret every LaTeX symbol as its ASCII code (as computers do), add 1 at the beginning of the resulting sequence of 0s and 1s, and interpret the resulting binary sequence as an integer in binary code.

Definition A2. *We say that a theory $\mathcal{M}$ is* stronger *than $\mathcal{L}$ if it contains all formulas, all axioms, and all deduction rules from $\mathcal{L}$, and also contains a special predicate* $\text{def}(n, x)$ *such that for every formula $P(x)$ from $\mathcal{L}$ with one free variable, the formula*

$$\forall y \, (\text{def}(\lfloor P(x) \rfloor, y) \leftrightarrow P(y))$$

is provable in $\mathcal{M}$.

The existence of a stronger theory can be easily proven: indeed, for $\mathcal{L}$=ZF, there exists a stronger theory $\mathcal{M}$. As an example of such a stronger theory, we can simply take the theory $\mathcal{L}$ plus all countably many equivalence formulas as described in Definition A2 (formulas corresponding to all possible formulas $P(x)$ with one free variable). This theory clearly contains $\mathcal{L}$ and all the desired equivalence formulas, so all we need to prove is that the resulting theory $\mathcal{M}$ is consistent (provided that $\mathcal{L}$ is consistent, of course). Due to compactness principle, it is sufficient to prove that for an arbitrary finite set of formulas $P_1(x), \ldots, P_m(x)$, the theory $\mathcal{L}$ is consistent with the above reflection-principle-type formulas corresponding to these properties $P_1(x), \ldots, P_m(x)$.

This auxiliary consistency follows from the fact that for such a finite set, we can take

$$\text{def}(n, y) \leftrightarrow (n = \lfloor P_1(x) \rfloor \,\&\, P_1(y)) \vee \ldots \vee (n = \lfloor P_m(x) \rfloor \,\&\, P_m(y)).$$

This formula is definable in $\mathcal{L}$ and satisfies all m equivalence properties. The statement is proven.

Important comments. In the main text, we will assume that a theory $\mathcal{M}$ that is stronger than $\mathcal{L}$ has been fixed; proofs will mean proofs in this selected theory $\mathcal{M}$.

An important feature of a stronger theory $\mathcal{M}$ is that the notion of an $\mathcal{L}$-definable set can be expressed within the theory $\mathcal{M}$: a set S is $\mathcal{L}$-definable if and only if

$$\exists n \in \mathbb{N} \, \forall y (\text{def}(n, y) \leftrightarrow y \in S).$$

In the paper, when we talk about definability, we will mean this property expressed in the theory $\mathcal{M}$. So, all the statements involving definability become statements from the theory $\mathcal{M}$ itself, *not* statements from metalanguage.