Logic of Scientific Discovery: How Physical Induction Affects What Is Computable and How It Is Related to Kolmogorov Complexity and Randomness

Vladik Kreinovich and Olga Kosheleva
University of Texas at El Paso
500 W. University, El Paso, Texas 79968, USA
vladik@utep.edu, olgak@utep.edu

Abstract: Most of our knowledge about a physical world comes from *physical induction*: if a hypothesis is confirmed by a sufficient number of observations, we conclude that this hypothesis is universally true. We show that a natural formalization of this property affects what is computable when processing measurement and observation results, and we explain how this formalization is related to Kolmogorov complexity and randomness. We also consider computational consequences of an alternative idea also coming form physics: that no physical law is absolutely true, that every physical law will sooner or later need to be corrected. It turns out that this alternative approach enables us to use measurement results go beyond what is usually computable.

Keywords: physical induction, computability, Kolmogorov complexity, randomness

1. Physical Induction

Physical induction: main idea. How do we come up with physical laws? Someone formulates a hypothesis, this hypothesis is tested, and if it confirmed sufficiently many times, then we conclude that this hypothesis is indeed a universal physical law. This conclusion is known as *physical induction*. Different physicists may disagree on how many experiments we need to become certain, but most physicists would agree that after sufficiently many confirmations, the hypothesis should be accepted as a physical law.

How to describe physical induction in precise terms. How can formalize physical induction? Let s denote the state of the world, and let $P(s, i)$ indicate that the property P holds in the i-th experiment. In these terms, physical induction means that for every property P, there is an integer N such that if the statements $P(s, 1)$, ..., $P(s, N)$ are all true, then the property P holds for all possible experiments – i.e., we have $\forall n\, P(s, n)$; see, e.g., (Kreinovich, 2012b).

Of course, this cannot be true for all *mathematically possible* states: it is theoretically possible that we have observation sequences in which an arbitrarily large number of "true" statements is followed by a "false" statement. Our understanding of the physicists' claims is that if we restrict ourselves to *physically meaningful* states, then physical induction is true. Thus, we arrive at the following definition.

Definition 1.1.

- *Let S be a set; its elements will be called* states of the world.
- *Let $T \subseteq S$ be a subset of S. We say that T* consists of physically meaningful states *if for every property P, there exists an integer N_P such that for each state $s \in T$ for which $P(s, i)$ holds for all $i \le N_P$, we have $\forall n\, P(s, n)$.*

Comments.

- To make this definition precise, we need to describe a language in which we can describe different properties. For this definition, we need to select a theory $\mathcal{L}$ which is rich enough to contain all physicists' arguments and at the same time weak enough so that we will be able to formally talk about definability in $\mathcal{L}$; for a detailed discussion, see Appendix (see also (Finkelstein & Kreinovich, 1987 ; Kreinovich, 2012b, 2009 ; Kreinovich & Finkelstein, 2004, 2006)).
- Instead of *properties*, we can reformulate this definition in terms of *definable sets*, i.e., sets of the type $\{x : P(x)\}$ corresponding to definable properties $P(x)$.

Definition 1.1′.

- *Let S be a set; its elements will be called* states of the world.
- *Let $T \subseteq S$ be a subset of S. We say that T* consists of physically meaningful states *if for every definable sequence of sets $\{A_n\}$, there exists an integer N_A such that*

$$T \cap \bigcap_{n=1}^{N_A} A_n = T \cap \bigcap_{n=1}^{\infty} A_n. \tag{1.1}$$

Existence proof. Do such subsets exist? The existence proof uses the fact that each definable sequence is described by a property in the corresponding language, and since there are no more than countably many words in a language, the set of all such properties is also (no more than) countable. For example, for the set S of real numbers or tuples, and for the case when all definable sets are measurable, we can enumerate all definable sequence, as $\{A_n^1\}$, $\{A_n^2\}$, ... Let us pick any number $\varepsilon \in (0,1)$. For each k, for the difference sets $D_n^k = \bigcap_{i=1}^{n} A_n^k - \bigcap_{i=1}^{\infty} A_n^k$, we have $D_{n+1}^k \subseteq D_n^k$ and $\bigcap_{n=1}^{\infty} D_n^k = \emptyset$, thus, $\mu(D_n^k) \to 0$. Hence, there exists n_k for which $\mu\left(D_{n_k}^k\right) \le 2^{-k}\cdot\varepsilon$. We then take $T = S - \bigcup_{k=1}^{\infty} D_{n_k}^k$. Here, $\mu\left(\bigcup_{k=1}^{\infty} D_{n_k}^k\right) \le \sum_{k=1}^{\infty} \mu\left(D_{n_k}^k\right) \le \sum_{k=1}^{\infty} 2^{-k}\cdot\varepsilon = \varepsilon < 1$, and thus, the difference T is non-empty. For this set T, we can take $N_{A^k} = n_k$.

From states of the world to specific quantities. In practice, we do not know the state of the entire world, we have a partial information about this state – e.g., we may know the values of finitely many quantities describing the current state of the world.

In precise terms, we can describe this by saying that we have a definable function $f : S \to X$ which maps every state of the world into the corresponding partial information – e.g., into the values of the corresponding quantities. One can prove that the range $f(T)$ corresponding to all physically meaningful states has the same property as the original set T:

Proposition 1.1. *Let S be a set, let a set $T \subseteq S$ consist of physically meaningful states, and let $f : S \to X$ be a definable function. Then, for every definable sequence of subsets $B_n \subseteq X$, there exists an integer N_B such that $f(T) \cap \bigcap_{n=1}^{N_B} B_n = f(T) \cap \bigcap_{n=1}^{\infty} B_n$.*

Proof. We want to prove that for some N_B, if an element $x \in f(T)$ belongs to the sets $B_1, \ldots, B_{N_B}$, then $x \in B_n$ for all n. An arbitrary element $x \in f(T)$ has the form $x = f(s)$ for some $s \in T$. Let us take $A_n \stackrel{\text{def}}{=} f^{-1}(B_n)$. Since T consists of physically meaningful states,

there exists an appropriate integer N_A. Let us take $N_B \stackrel{\text{def}}{=} N_A$. By definition of A_n, the condition $x = f(s) \in B_i$ implies that $s \in A_i$; so we have $s \in A_i$ for all $i \le N_A$. Thus, we have $s \in A_n$ for all n, which implies that $x = f(s) \in B_n$. The proposition is proven.

Comment. Because of this proposition, we can view the set X as the set of states of the world.

If we limit ourselves to physically meaningful states, then many generally-undecidable computational problems become decidable. It is known that many computational problems are, in general, algorithmically undecidable.

For example, from the physical viewpoint, real numbers x describe values of different quantities. We get values of real numbers by measurements; measurements are never 100% accurate, so what we get after each measurement is an approximate value r_k of x. In principle, we can measure x with higher and higher accuracy. Thus, from the computational viewpoint, the information that we have about a number x is a sequence of rational numbers r_k for which, e.g., $|x - r_k| \le 2^{-k}$. By an algorithm processing real numbers, it is this reasonable to understand an algorithm which is allowed to send a request k and get the corresponding rational value r_k. In other words, the algorithm can use the sequence r_k as an "oracle' (subroutine). This is how computations with real numbers are defined in *computable analysis*; see, e.g., (Pour-El & Richards, 1989 ; Weihrauch, 2000).

It is known (Pour-El & Richards, 1989 ; Weihrauch, 2000) that equality of real numbers is undecidable: no algorithm is possible which, given two real numbers x and y, checks whether $x = y$. For physically meaningful real numbers, however, a deciding algorithm *is* possible.

Proposition 1.2. (Kreinovich, 2012a) *For every set $T \subseteq \mathbb{R}^2$ which consists of physically meaningful pairs (x, y) of real numbers, there exists an algorithm which decides whether $x = y$.*

Proof. We can take $A_n = \{(x, y) : 0 < |x - y| < 2^{-n}\}$. Then, there exists an integer N_A for which the property (1.1) holds. In this case, the intersection of all the sets A_n is empty, thus (1.1) means that T has no elements from the from the intersection $\bigcap\limits_{n=1}^{N_A} A_n$ – which coincides with A_{N_A}. Thus, for each pair $(x, y) \in T$, we have either $x = y$ or $|x - y| \ge 2^{-N_A}$. If we take $2^{-(N_A+3)}$-approximations x' and y' to x and y, then:

- in the case $x = y$ we will get $|x' - y'| \le (1/4) \cdot 2^{-N_A}$, and
- in the case $|x - y| \ge 2^{-N_A}$, we get $|x' - y'| \ge (3/4) \cdot 2^{-N_A}$.

These cases are clearly distinguishable. The proposition is proven.

Another known negative result is that in general, it is not possible, given a function $f(x)$ which attains negative and positive values, to compute its root (see (Kreinovich, 2012a ; Pour-El & Richards, 1989 ; Weihrauch, 2000) for exact definitions). This becomes possible if we restrict ourselves to physically meaningful functions:

Proposition 1.3. *Let K be a computable compact, and let X be the set of all functions*

$$f : K \to \mathbb{R}$$

that attains 0 value somewhere on K. Then, for every set $T \subseteq X$ consisting of physically meaningful functions, there exists an algorithm that, given a function $f \in T$, computes a point x at which $f(x) = 0$.

Moreover, we can not only produce a root x, we can actually compute, for any given n, an 2^{-n}-approximations to the corresponding set of roots $\{x : f(x) = 0\}$ in the sense of the Hausdorff distance

$$d_H(A, B) \stackrel{\text{def}}{=} \max\left(\sup_{a\in A} d(a, B), \sup_{b\in B} d(b, A)\right),$$

where $d(a, B) \stackrel{\text{def}}{=} \inf\limits_{b\in B} d(a, b)$.

In other words, there exists an algorithm that, given a physically meaningful function $f(x)$ on a computable compact K that attains a 0 value somewhere on K, computes a root x – and also computes an 2^{-n}-approximations to the corresponding set of roots.

Proof. To compute the set $R \stackrel{\text{def}}{=} \{x : f(x) = 0\}$ with accuracy $\varepsilon > 0$, let us take an $(\varepsilon/2)$-net $\{x_1, \ldots, x_n\} \subseteq K$. Such a net exists, by definition of a computable compact set.

For each i, we can compute $\varepsilon' \in (\varepsilon/2, \varepsilon)$ for which $B_i \stackrel{\text{def}}{=} \{x : d(x, x_i) \le \varepsilon'\}$ is a computable compact set; see, e.g., (Bishop, 1967). It is possible to algorithmically compute the maximum of a function on a computable compact set; thus, we can compute the value

$$m_i \stackrel{\text{def}}{=} \min\{|f(x)| : x \in B_i\}.$$

Since $f \in T$, similarly to the proof of Proposition 1.2, we can prove that there exists an integer N for which, for all $f \in T$ and for all i, we have either $m_i = 0$ or $m_i \ge 2^{-N}$. Thus, by computing each m_i with accuracy $2^{-(N+2)}$, we can check whether $m_i = 0$ or $m_i > 0$.

We claim that $d_H(R, \{x_i : m_i = 0\}) \le \varepsilon$, i.e., that:

- for every point x_i for which $m_i = 0$, there exists an ε-close root x, and
- for every root x, there exists an ε-close point x_i for which $m_i = 0$.

Indeed, if $m_i = 0$, this means that the minimum of a function $|f(x)|$ on the ε'-ball B_i with a center in x_i is equal to 0. Since the set K is compact, this value 0 is attained, i.e., there exists a value $x \in B_i$ for which $f(x) = 0$. From $x \in B_i$, we conclude that $d(x, x_i) \le \varepsilon'$ and, since $\varepsilon' < \varepsilon$, that $d(x, x_i) < \varepsilon$. Thus, x_i is ε-close to the root x.

Vice versa, let x be a root, i.e., let $f(x) = 0$. Since the points x_i form an $(\varepsilon/2)$-net, there exists an index i for which $d(x, x_i) \le \varepsilon/2$. Since $\varepsilon/2 < \varepsilon'$, this means that $d(x, x_i) \le \varepsilon'$ and thus, $x \in B_i$. Therefore, $m_i = \min\{|f(x)| : x \in B_i\} = 0$. So, the root x is ε-close to a point x_i for which $m_i = 0$. The proposition is proven.

Similar results hold for two other problems which are, in general, not algorithmically solvable: locating the maxima (or minima) and finding the fixed points:

Proposition 1.4. *Let K be a computable compact, and let X be the set of all functions*

$$f : K \to \mathbb{IR}.$$

Then, for every set $T \subseteq X$ consisting of physically meaningful functions, there exists an algorithm that, given a function $f \in T$, computes a point x at which $f(x) = \max\limits_{y} f(y)$.

Moreover, we can not only produce such a point, we can actually compute, for any given n, an 2^{-n}-approximations to the corresponding set of global maximum locations $\left\{x : f(x) = \max\limits_{y} f(y)\right\}$.

Proof. This problem can be reduced to the previous one if we take into consideration the fact that maximum $\max\limits_y f(y)$ of a computable function on a computable compact is computable and that that $f(x) = \max\limits_y f(y)$ if and only if $g(x) = 0$, where $g(x) \stackrel{\text{def}}{=} f(x) - \max\limits_y f(y)$.

Proposition 1.5. *Let K be a computable compact, and let X be the set of all functions*

$$f : K \to K$$

that have a fixed point x for which $f(x) = x$. Then, for every set $T \subseteq X$ consisting of physically meaningful functions, there exists an algorithm that, given a function $f \in \mathcal{T}$, computes a point x at which $f(x) = x$.

Moreover, we can not only produce such a fixed point, we can actually compute, for any given n, an 2^{-n}-approximation to the corresponding set of all fixed points $\{x : f(x) = x\}$.

Proof. This problem can be reduced to the root finding problem if we take into consideration that that $f(x) = x$ if and only if $g(x) = 0$, where $g(x) \stackrel{\text{def}}{=} d(f(x), x)$.

Relation to randomness and Kolmogorov complexity. Kolmogorov and Martin-Löf proposed a new definition of a random sequence (Li & Vitányi, 2008), a definition that separates physically random binary sequences – e.g., sequences that appear in coin flipping experiments or sequences that appear in quantum measurements – from sequence that follow some pattern. Intuitively, if a sequence s is random, it satisfies all the probability laws – like the law of large numbers, the central limit theorem, etc. Vice versa, if a sequence satisfies all probability laws, then for all practical purposes we can consider it random. Thus, we can define a sequence to be random if it satisfies all probability laws.

What is a probability law? In precise terms, it is a statement S which 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.

Equivalently, this statement can be reformulated as follows: a sequence s is *not* random if $s \in C$ for a (definable) set C ($= -S$) with $P(C) = 0$. As a result, we arrive at the following definition: *We say that a sequence is* random *if it does not belong to any definable set of measure 0.*

If we use different languages to formalize the notion "definable", we get different versions of Kolmogorov-Martin-Löf randomness. In particular, if we define "definable" properties as computable ones (in some reasonable sense), then randomness can be described in terms of *Kolmogorov complexity* $K(x)$ of a string – which is defined as the shortest (bit) length of a program (in a fixed universal programming language) which generates x. Crudely speaking, an infinite string $s = s_1 s_2 \ldots$ is random if, for some constant $C > 0$, we have $K(s_1 \ldots s_n) \geq n - C$ (see Section 2 for more details).

This definition makes sense – each program uniquely determines its outcome, so to generate 2^n different sequences of length n, we need to have 2^n different programs – and since for $C \gg 0$, there are only $2^{n-C} \ll 2^n$ possible binary programs, most sequences require programs of length $K \geq n - C$. A non-trivial part of this definition is that this inequality is sufficient to guarantee that all other properties are satisfied (Li & Vitányi, 2008).

The above definition means that (definable) events with probability 0 cannot happen. In practice, physicists also assume that events with a *very small* probability cannot happen. For example, they believe that it is not possible that all the molecules in the originally uniform air move to one side of the room – although, from the viewpoint of statistical physics, the probability of this event is not zero. This fits very well with a commonsense understanding of

rare events: e.g., if a coin falls head 100 times in a row (or a casino roulette gets to red 100 times in a row), any reasonable person will conclude that this coin is not fair.

It is not possible to formalize this idea by simply setting a threshold $p_0 > 0$ below which events are not possible – since then, for N for which $2^{-N} < p_0$, no sequence of N heads or tails would be possible at all. However, we know that for each decreasing ($A_n \supseteq A_{n+1}$) sequence of properties A_n with $\lim p(A_n) = 0$ (e.g., A_n = "we can get first n heads"), there exists an N above which a truly random sequence cannot belong to A_N. In (Finkelstein & Kreinovich, 1987 ; Kreinovich, 2012a, 2012b, 2009 ; Kreinovich & Finkelstein, 2004, 2006), we thus propose the following definition of a set of random elements:

Definition 1.2. *We say that $\mathcal{R}$ is a* set of random elements *if for every definable decreasing sequence $\{A_n\}$ for which $\lim P(A_n) = 0$, there exists an N for which*

$$\mathcal{R} \cap A_N = \emptyset.$$

Let $\mathcal{R}_K$ denote the set of the elements random in the usual Kolmogorov-Martin-Löf sense. Then the following is true (Kreinovich, 2009, 2012a):

Proposition 1.6.

- *Every set of random elements consists of physically meaningful elements. elements.*
- *For every set T consisting of physically meaningful elements, the intersection $T \cap \mathcal{R}_K$ is a set of random elements.*

Proof. For every definable sequence A_n, the corresponding difference sets $D_n \stackrel{\text{def}}{=} \bigcap_{i=1}^{n} A_i - \bigcap_{i=1}^{\infty} A_i$ are decreasing ($D_n \supseteq D_{n+1}$) and their intersection is empty. Thus, $P(D_n) \to 0$. Therefore, there exists an N for which the set of random elements does not contain any elements from D_N. As we have shown in the existence proof, this implies that for this N, the property (1.1) is satisfied. Thus, every set of random elements indeed consists of physically meaningful elements.

Vice versa, let T be a set consisting of physically meaningful elements. Let us prove that $\mathcal{T} \cap \mathcal{R}_K$ is a set of random elements. Indeed, if $A_n \supseteq A_{n+1}$ and $P\left(\bigcap_{n=1}^{\infty} A_n\right) = 0$, then for $B_m \stackrel{\text{def}}{=} A_m - \bigcap_{n=1}^{\infty} A_n$, we have $B_m \supseteq B_{m+1}$ and $\bigcap_{n=1}^{\infty} B_n = \emptyset$. Thus, by definition of a set consisting of physically meaningful elements, we conclude that the property (1.1) holds, i.e., in our case, that $B_N \cap T = \emptyset$. Since $P\left(\bigcap_{n=1}^{\infty} A_n\right) = 0$, we also know that $\left(\bigcap_{n=1}^{\infty} A_n\right) \cap \mathcal{R}_K = \emptyset$. Thus, $A_N = B_N \cup \left(\bigcap_{n=1}^{\infty} A_n\right)$ has no common elements with the intersection $T \cap \mathcal{R}_K$. The proposition is proven.

2. Alternative Idea: No Physical Theory Is Perfect

No physical theory is perfect: a common physicists' belief. Physical induction implies that a physical law is universally valid. However, 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., (Feynman, Leighton, & Sands, 2005).

How does this belief affect computations? In this section, following (Kosheleva & Soloviev, 1981 ; Zakharevich & Kosheleva, 2014), we show how this idea affects computations.

Towards formalizing the above belief. 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.

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? 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.

Of course, 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.

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 ($\lim \omega^{(n)} = \omega$), 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 $\omega_1\ldots\omega_n$ 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*.

The belief 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 $\omega_1 \ldots \omega_n$ 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 $\omega_1 \ldots \omega_n$ 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 definitions. By combining the above properties of a set T which describes a physical theory, we arrive at the following definition.

Definition 2.1. *By a* physical theory*, we mean a non-empty closed nowhere dense definable set T.*

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 2.2. *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 2.1).*

Comment. Do such sequences exist at all? Yes, they do. 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., (Jalal-Kamali, Nebesky, Durcholz, Kreinovich, & Longpré, 2012 ; Oxtoby, 1980).

How to describe general computations. We want to analyze how the physicists' belief affects computations. To be able to perform this analysis, we need to describe what exactly we mean by 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 $\alpha = \alpha_1 \ldots \alpha_n \ldots$ 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 A. N. 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., (Li & Vitányi, 2008).

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.

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

Proposition 2.1. *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 proven.

Acknowledgments. This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721.

The authors are thankful to Luc Longpré, Sergei Soloviev, and Michael Zakharevich for valuable discussions.

Références

Bishop, E. (1967). *Foundations of Constructive Analysis*. New York : McGraw-Hill.

Feynman, R., Leighton, R., & Sands, M. (2005). *The Feynman Lectures on Physics*. Reading, MA : Addison Wesley.

Finkelstein, A. M., & Kreinovich, V. (1987). Impossibility of hardly possible events: physical consequences. In *Abstracts of the 8th international congress on logic, methodology and philosophy of science* (Vol. 5(2), pp. 25–27). Moscow.

Jalal-Kamali, A., Nebesky, O., Durcholz, M. H., Kreinovich, V., & Longpré, L. (2012). Towards a "generic" notion of genericity: from 'typical' and 'random' to meager, shy, etc. *Journal of Uncertain Systems*, *6(2)*, 104–113.

Kosheleva, O. M., & Soloviev, S. V. (1981). On the logic of using observable events in decision making. In *Proceedings of the IX National USSR Symposium on Cybernetics* (pp. 49–51). Moscow.

Kreinovich, V. (2009). Toward formalizing non-monotonic reasoning in physics: the use of kolmogorov complexity. *Revista Iberoamericana de Inteligencia Artificial*, *41*, 4–20.

Kreinovich, V. (2012a). Negative results of computable analysis disappear if we restrict ourselves to random (or, more generally, typical) inputs. *Mathematical Structures and Modeling*, *25*, 100–113.

Kreinovich, V. (2012b). Towards formalizing non-monotonic reasoning in physics: logical approach based on physical induction and its relation to Kolmogorov complexity. In E. Erdem, J. Lee, Y. Lierler, & D. Pearce (Eds.), *Correct Reasoning: Essays on Logic-Based AI in Honor of Vladimir Lifschitz* (Vol. 7265, pp. 390–404). Berlin, Heidelberg, New York : Springer Verlag.

Kreinovich, V., & Finkelstein, A. M. (2004). Towards applying computational complexity to foundations of physics. *Notes of Mathematical Seminars of St. Petersburg Department of Steklov Institute of Mathematics*, *316*, 63–110.

Kreinovich, V., & Finkelstein, A. M. (2006). Towards applying computational complexity to foundations of physics. *Journal of Mathematical Sciences*, *134*, 2358–2382.

Li, M., & Vitányi, P. M. B. (2008). *An Introduction to Kolmogorov Complexity*. Berlin, Heidelberg, New York : Springer-Verlag.

Oxtoby, J. C. (1980). *Measure and Category: A Survey of the Analogies Between Topological and Measure Spaces*. New York, Heidelberg, Berlin : Springer Verlag.

Pour-El, M. B., & Richards, J. I. (1989). *Computability in Analysis and Physics*. Berlin : Springer Verlag.

Weihrauch, K. (2000). *Computable Analysis*. Berlin : Springer Verlag.

Zakharevich, M., & Kosheleva, O. (2014). If many physicists are right and no physical theory is perfect, then the use of physical observations can enhance computations. *Journal of Uncertain Systems*, *8*, to appear.

Appendix A: A Formal Definition of Definable Sets

Definition A.1. *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 $\{x \mid P(x)\}$ is defined in the theory $\mathcal{L}$. We will then call the set $\{x \mid P(x)\}$* $\mathcal{L}$-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 A.2. *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 $\mathrm{def}(n, x)$ such that for*

every formula $P(x)$ from $\mathcal{L}$ with one free variable, the formula

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

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

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

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

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

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.