Toward Formalizing Non-Monotonic Reasoning in Physics: the Use of Kolmogorov Complexity and Algorithmic Information Theory to Formalize the Notions "Typically" and "Normally"

Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso,
El Paso, TX 79968, USA, vladik@cs.utep.edu

Abstract

When a physicist writes down equations, or formulates a theory in any other terms, he usually means not only that these equations are true for the real world, but also that the model corresponding to the real world is "typical" among all the solutions of these equations. This type of argument is used when physicists conclude that some property is true by showing that it is true for "almost all" cases. There are formalisms that partially capture this type of reasoning, e.g., techniques based on the Kolmogorov-Martin-Lof definition of a random sequence. The existing formalisms, however, have difficulty formalizing, e.g., the standard physicists' argument that a kettle on a cold stove cannot start boiling by itself, because the probability of this event is too small.

We present a new formalism that can formalize this type of reasoning. This formalism also explains "physical induction" (if some property is true in sufficiently many cases, then it is true), and many other types of physical reasoning.

1 Preface: Is the Current Mathematical Formalization of Physics Fully Adequate?

Problem that Led to the Notion of Kolmogorov Complexity. Traditional mathematical approach to the analysis of physical systems implicitly assumed that all mathematically possible integers are physically possible as well, and all mathematically possible trajectories are physically possible. Traditionally, this approach has worked well in physics and in engineering, but it does not lead to a very good understanding of chaotic systems, which, as is now known, are extremely important in the study of real-world phenomena ranging from weather to biological systems.

Kolmogorov was among the first who started, in the 1960s, analyzing the discrepancy between the physical and the mathematical possibility. He pinpointed two main reasons why a mathematical correct solution to the corresponding system of differential or difference equation can be not physically possible:

First, there is a difference in understanding the term "random" in mathematics and in physics. For example, in statistical physics, it is possible (probability is positive) that a kettle, when placed on a cold stove, will start boiling by itself. From the viewpoint of a working physicist, however, this is absolutely impossible. Similarly, a trajectory which requires a highly unprobable combination of initial conditions may be mathematically correct, but, from the physical viewpoint, it is impossible.

Second, the traditional mathematical analysis tacitly assumes that all integers and all real numbers, no matter how large or how small, are physically possible. From the engineering viewpoint, however, a number like $10^{10^{10}}$ is not possible at all, because it exceeds the number of particles in the Universe. In particular, solutions to the corresponding systems of differential equations which lead to some numbers may be mathematically correct, but they are physically meaningless.

Attempts to formalize these restrictions have been started by Kolmogorov himself. These attempts are at present, mainly undertaken by researchers in theoretical computer science who face a similar problem of distinguishing between theoretically possible "algorithms" and feasible practical algorithms which can provide the results of their computations in reasonable time.

The goal of the present paper is to use the experience of formalizing these notions in theoretical computer science to enhance the formalization of similar constraints in engineering and physics.

What Is Kolmogorov Complexity. This research is mainly concentrated around the notion of Kolmogorov complexity. This notion was introduced independently by several people: Kolmogorov in Russia and Solomonoff and Chaitin in the US. Kolmogorov used it to formalize the notion of a random sequence. Probability theory describes most of the physicist intuition in precise mathematical terms, but it does not allow us to tell whether a given finite sequence of 0's and 1's is random or not. Kolmogorov defined a complexity $K(x)$ of a binary sequence x as the shortest length of a program which produces this sequence. Thus, a sequence consisting of all 0's or a sequence 010101... have a very short Kolmogorov complexity because these sequences can be generated by simple programs, while for a sequence of results of tossing a coin, probably the shortest program is to write print(0101...) and then reproduce the entire sequence. Thus, when $K(x)$ is approximately equal to the length $\mathrm{len}(x)$ of a sequence, this sequence is random, otherwise it is not. The best source for Kolmogorov complexity is a book [13].

Beyond Traditional Kolmogorov Complexity. The definition of $K(x)$ only takes into consideration the length $\mathrm{len}(p)$ of a program p. From the physical viewpoint, it is also important to take into consideration its running time $t(p)$, because if it exceeds the lifetime of the Universe, this algorithm makes no practical sense. This development is in line with Kolmogorov's original idea that some natural numbers which are mathematically possible (like $10^{10^{10}}$) are not feasible and thus, should not considered as feasible. Corresponding modifications are also described in the above book. We plan to show how to use the corresponding ideas in physics and engineering.

Specifically, these ideas lead us to the following improvements in comparison with the traditional mathematical approaches to science and engineering, approaches that do not take into consideration the difference between "inhuman" ("abnormal") and "human" ("normal") numbers:

Physically impossible events become "mathematically impossible" as well. From the physical and engineering viewpoints, a cold kettle placed on a cold stove will never start boiling by itself. However, from the traditional probabilistic viewpoint, there is a positive probability that it will. Our new approach makes the mathematical formalism consistent with common sense: crudely speaking, the probability is so small that this event is simply physically impossible.

Physically possible indirect measurements become "mathematically possible" as well. In engineering and in physics, we often cannot directly measure the desired quantity; instead, we measure related properties and then use the measurement results to reconstruct the measured values. In mathematical terms, the corresponding reconstruction problem is called the *inverse problem*. In practice, this problem is efficiently used to reconstruct the signal from noise, to find the faults within a metal plate, etc. However, from the purely mathematical viewpoint, most inverse problems are *ill-defined* meaning that we cannot really reconstruct the desired values without making some additional assumptions.

What we are planning to do. In this paper, we show that the only assumption we need to make is that the reconstructed signal, etc., is "normal", and immediately, the problem becomes well-defined in the precise mathematical sense. We also show that this idea naturally leads to an emergence of *chaos*, and it also helps to deal with systems that display chaotic behavior.

2 Physicists Assume that Initial Conditions and Values of Parameters Are Not Abnormal

To a mathematician, the main contents of a physical theory is the equations. The fact that the theory is formulated in terms of well-defined mathematical equations means that the actual field must satisfy these equations. However, this fact does *not* mean that *every* solution of these equations has a physical sense. Let us give three examples:

Example 1. At any temperature greater than absolute zero, particles are randomly moving. It is theoretically possible that all the particles start moving in one direction, and, as a result, the chair that I am sitting on starts lifting up into the air. The probability of this event is small (but positive), so, from the purely mathematical viewpoint, we can say that this event is possible but highly unprobable. However, the physicists say plainly that such an abnormal event is *impossible* (see, e.g., [5]).

Example 2. Another example from statistical physics: Suppose that we have a two-chamber camera. The left chamber if empty, the right one has gas in it. If we open the door between the chambers, then the gas would spread evenly between the two chambers. It is theoretically possible (under appropriately chosen initial conditions) that the gas that was initially evenly distributed would concentrate in one camera, but physicists believe this abnormal event to be impossible. This is a general example of what physicists call *irreversible processes*: on the atomic level, all equations are invariant with respect to changing the order of time flow $t \to -t$). So, if we have a process that goes from state A to state B, then, if at B, we revert all the velocities of all the atoms, we will get a process that goes from B to A. However, in real life, many processes are clearly irreversible: an explosion can shatter a statue, but it is hard to imagine an inverse process: an implosion that glues together shattered pieces into a statue. Boltzmann himself, the 19 century author of statistical physics, explicitly stated that such inverse processes "may be regarded as impossible, even though from the viewpoint of probability theory that outcome is only extremely improbable, not impossible." [1].

Example 3. If we flip a coin 100 times in a row, and get heads all the time, then a person who is knowledgeable in probability would say that it is possible, while an engineer (and any person who uses common sense reasoning) would say that the coin is not fair, because if it is was a fair coin, then this abnormal event would be impossible.

In all the above cases, we knew something about probability. However, there are examples of this type of reasoning in which probability does not enter into picture at all. For example, in general relativity, it is known that for almost all initial conditions (in some reasonable sense) the solution has a singularity point. Form this, physicists conclude that the solution that corresponds to the geometry of the actual world has a singularity (see, e.g., [14]): the reason is that the initial conditions that lead to a non-singularity solution are abnormal (atypical), and the actual initial conditions must be not abnormal.

In all these cases, the physicists (implicitly or explicitly) require that the actual values of the fields must not satisfy the equations, but they must also satisfy the additional condition: that the initial conditions should *not* be *abnormal*.

Comment: why is it non-monotonic? Since we are presenting this text at the meeting devoted to studying non-monotonic reasoning, readers who are not familiar with this type of reasoning in phsyics may ask: is this reasoning really non-monotonic? Let us give two main reasons why it is.

First, we are trying to formalize the notion of "typical", with the corresponding implication relation "typically, A implies B" which is, in this case, interpreted as "B is true in all not abnormal situations in which A is true".

Second, as we will discuss in the next section, often, this "typicalness" can be described in probabilistic terms: "typically, A implies B" means that out of all events for which A is true, the probability of "not B" is smaller than a certain threshold p_0. Even in this case, such an "implication" does not correspond to a standard monotonic logic: e.g., it is not transitive: if $C \subseteq B \subseteq A$ with $p(A) = 1$, $p(B) = 1 - p_0$, and

$p(C) = (1 - p_0)^2$, then "typically, A implies B" and "typically, B implies C", but *not* "typically, A implies C".

3 The Notion of "Not Abnormal" Is Difficult to Formalize

At first glance, it looks like in the probabilistic case, this property has a natural formalization: if a probability of an event is small enough (say, $\leq p_0$ for some very small p_0), then this event cannot happen. For example, the probability that a fair coin falls heads 100 times in a row is 2^{-100}, so, if we choose $p_0 \geq 2^{-100}$, then we will be able to conclude that such an event is impossible. The problem with this approach is that *every* sequence of heads an details has exactly the same probability. So, if we choose $p_0 \geq 2^{-100}$, we will thus exclude all possible sequences of heads and tails as physically impossible. However, anyone can toss a coin 100 times, and this prove that some sequences are physically possible.

Historical comment. This problem was first noticed by Kyburg under the name of *Lottery paradox* [12]: in a big (e.g., state-wide) lottery, the probability of winning the Grand Prize is so small, then a reasonable person should not expect it. However, some people do win big prizes.

4 How to Formalize the Notion of "Not Abnormal": Idea

"Abnormal" means something unusual, rarely happening: if something is rare enough, it is not typical ("abnormal"). Let us describe what, e.g., an abnormal height may mean. If a person's height is ≥ 6 ft, it is still normal (although it may be considered abnormal in some parts of the world). Now, if instead of 6 pt, we consider 6 ft 1 in, 6 ft 2 in, etc, then sooner or later we will end up with a height h such that everyone who is higher than h will be definitely called a person of abnormal height. We may not be sure what exactly value h experts will call "abnormal", but we are sure that such a value exists.

Let us express this idea is general terms. We have a *Universe of discourse*, i.e., a set U of all objects that we will consider. Some of the elements of the set U are abnormal (in some sense), and some are not. Let us denote the set of all elements that are *typical* (*not abnormal*) by T. On this set, we have a decreasing sequence of sets $A_1 \supseteq A_2 \supseteq \ldots \supseteq A_n \supseteq \ldots$ with the property that $\cap A_n = \emptyset$. In the above example, U is the set of all people, A_1 is the set of all people whose height is ≥ 6 ft, A_2 is the set of all people whose height is ≥ 6 ft 1 in, A_2 is the set of all people whose height is ≥ 6 ft 2 in, etc. We know that if we take a sufficiently large n, then all elements of A_n are abnormal (i.e., none of them belongs to the set T of not abnormal elements). In mathematical terms, this means that for some n, we have $A_n \cap T = \emptyset$.

In case of a coin: U is the set of all infinite sequences of results of flipping a coin; A_n is the set of all sequences that start with n heads but have some tail afterwards. Here, $\cup A_n = \emptyset$. Therefore, we can conclude that there exists an n for which all elements of A_n are abnormal. So, if we assume that in our world, only not abnormal initial conditions can happen, we can conclude that for some n, the actual sequence of results of flipping a coin cannot belong to A_n. The set A_n consists of all elements that start with n heads and a tail after that. So, the fact that the actual sequence does not belong to A_n means that if an actual sequence has n heads, then it will consist of all heads. In plain words, if we have flipped a coin n times, and the results are n heads, then this coin is biased: it will always fall on heads.

Let us describe this idea in mathematical terms [7, 11]. To make formal definitions, we must fix a formal theory: e.g., the set theory ZF (the definitions and results will not depend on what exactly theory we choose). A set S is called *definable* if there exists a formula $P(x)$ with one (free) variable x such that $P(x)$ if and only if $x \in S$.

Crudely speaking, a set is definable if we can *define* it in ZF. The set of all real numbers, the set of all solutions of a well-defined equations, every set that we can describe in mathematical terms is definable.

This does not means, however, that *every* set is definable: indeed, every 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 definable sets. Since, e.g., there are more than denumerably many set of integers, some of them are thus not definable.

Definition 1 *A sequence of sets $A_1, \ldots, A_n, \ldots$ is called* definable *if there exists a formula $P(n, x)$ such that $x \in A_n$ if and only if $P(n, x)$.*

Definition 2 *Let U be a universal set.*

- *A non-empty set $T \subseteq U$ is called a* set of typical (not abnormal) elements *if for every definable sequence of sets A_n for which $A_n \supseteq A_{n+1}$ and $\cap A_n = \emptyset$, there exists an N for which $A_N \cap T = \emptyset$.*
- *If $u \in T$, we will say that u is* not abnormal.
- *For every property P, we say that* "normally, for all u, $P(u)$" *if $P(u)$ is true for all $u \in T$.*

It is possible to prove that abnormal elements do exist [7]; moreover, we can select T for which abnormal elements are as rare as we want: for every probability distribution p on the set U and for every ε, there exists a set T for which the probability $p(x \notin T)$ of an element to be abnormal is $\leq \varepsilon$:

Proposition 1 *For every probability measure μ on a set U (in which all definable sets are measurable), and for every $\varepsilon > 0$, there exists a set T of typical elements for which $\mu(T) > 1 - \varepsilon$.*

Proof. Similarly to the above argument, one can show that there are no more than countably many definable sequences of sets $\{A_n\}$. Thus, there are at most countably many definable decreasing sequences $a = \{A_n\}$ for which $\cap A_n = \emptyset$. Therefore, we can order all such sequences into a sequence of sequences: $a^{(1)} = \{A_n^{(1)}\}$, $a^{(2)} = \{A_n^{(2)}\}$, ... For each of these sequences $a^{(k)}$, since $\cap A_n^{(k)} = \emptyset$, we have $\mu(A_n^{(k)}) \to 0$ as $n \to \infty$, hence there exists an N_k for which $\mu(A_{N_k}^{(k)}) < \varepsilon/2^k$.

Let us show that as T, we can take the complement $U \setminus A$ to the union A of all the sets $A_{N_k}^{(k)}$. Indeed, by our choice of T, for every definable decreasing sequence $a^{(k)} = \{A_n^{(k)}\}$, there exists an N, namely $N = N_k$, for which $T \cap A_N^{(k)} = \emptyset$.

To complete the proof, we must show that $\mu(T) > 1 - \varepsilon$. Indeed, from $\mu(A_{N_k}^{(k)}) < \varepsilon/2^k$, we conclude that $\mu(A) = \mu(\cup A_{N_k}^{(k)}) \leq \sum \mu(A_{N_k}^{(k)}) < \sum \varepsilon/2^k = \varepsilon$, and therefore, $\mu(T) = \mu(U \setminus A) = 1 - \mu(A) > 1 - \varepsilon$.

5 Application: Restriction to "Not Abnormal" Solutions Leads to Regularization of Ill-Posed Problems

An ill-posed problem arises when we want to reconstruct the state s from the measurement results r. Usually, all physical dependencies are continuous, so, small changes of the state s result in small changes in r. In other words, a mapping $f : S \to R$ from the set of all states to the set of all observations is continuous (in some natural topology). We consider the case when the measurement results are (in principle) sufficient to reconstruct s, i.e., the case when the mapping f is 1-1. That the problem is ill-posed means that small changes in r can lead to huge changes in s, i.e., that the inverse mapping $f^{-1} : R \to S$ is *not* continuous.

We will show that if we restrict ourselves to states S that are not abnormal, then the restriction of f^{-1} will be continuous, and the problem will become well-posed.

Definition 3 *A definable metric space (X, d) is called* definably separable *if there exists a definable everywhere dense sequence $x_n \in X$.*

Proposition 2 *Let S be a definably separable definable metric space, T be a set of all not abnormal elements of S, and $f : S \to R$ be a continuous 1-1 function. Then, the inverse mapping $f^{-1} : R \to S$ is continuous for every $r \in f(T)$.*

In other words, if we know that we have observed a not abnormal state (i.e., that $r = f(s)$ for some $s \in T$), then the reconstruction problem becomes well-posed. So, if the observations are accurate enough, we get as small guaranteed intervals for the reconstructed state s as we want.

Proof. It is known that if a set K is compact, then for any 1-1 continuous function $K \to R$, its inverse is also continuous. Thus, to prove our result, we will show that the closure $\overline{T}$ of the set T is compact.

A set K in a metric space S is compact if and only it is closed, and for every positive real number $\varepsilon > 0$, it has a finite ε-net, i.e., a finite set $K(\varepsilon)$ with the property that every $s \in K$, there exists an element $s(\varepsilon) \in K(\varepsilon)$ that is ε-close to s.

The closure $K = \overline{T}$ is clearly closed, so, to prove that this closure is compact, it is sufficient to prove that it has a finite ε-set for all $\varepsilon > 0$. For that, it is sufficient to prove that for every $\varepsilon > 0$, there exists a finite ε-net for the set R.

If a set T has a ε-net $T(\varepsilon)$, and $\varepsilon' > \varepsilon$, then, as one can easily see, this same set $T(\varepsilon)$ is also a ε'-net for T. Therefore, it is sufficient to show that finite ε-nets for T exist for $\varepsilon = 2^{-k}, k = 0, 1, 2, \ldots$

Let us fix $\varepsilon = 2^{-k}$. Since the set S is definably separable, there exists a definable sequence $x_1, \ldots, x_i, \ldots$ which is everywhere dense in S. As A_n, we will now take the complement to the union U_n of n closed balls $B_\varepsilon(x_1), \ldots, B_\varepsilon(x_n)$ of radius ε with centers in $x_1, \ldots, x_n$.

Clearly, $A_n \supseteq A_{n+1}$. Since x_i is an everywhere dense sequence, for every $s \in S$, there exists an n for which $s \in B_\varepsilon(x_n)$ and for which, therefore, $s \in U_n$ and $x \notin A_n = S \setminus U_n$. Hence, the intersection of all the sets A_n is empty.

Therefore, according to the definition of a set of typical elements, there exists an N for which $T \cap A_N = \emptyset$. This means that $T \subseteq U_N$. This, in its turn, means that the elements $x_1, \ldots, x_N$ form an ε-net for T. So, the set T has a finite ε-net for $\varepsilon = 2^{-k}$.

Comment. To actually use this result, we need an *expert* who will tell us what is abnormal.

6 Application: Every Physical Quantity is Bounded

Proposition 3 *If U is a definable set, and $f : U \to R$ is a definable function, then there exists a number C such that if $u \in U$ is not abnormal, then $|f(u)| \le C$.*

Proof. We can take $A_n \stackrel{\text{def}}{=} \{u \mid |f(u)| > n\}$; then, $\cap A_n = \emptyset$, hence there exists N for which $A_N \cap T = \emptyset$, i.e., for which, once $u \in T$, we have $u \notin A_N$ – i.e., $|f(u)| \le N$.

If we use the physicists' idea that abnormal initial conditions and/or abnormal values of parameters are impossible, then we can make the following conclusions:

Special relativity. If as U, we take the set of all the particles, and as f, we take velocity, then we can conclude that the velocities of all (not abnormal) particles is bounded by some constant C. This is exactly what special relativity says, with the speed of light as C.

Cosmology. If we take the same state U, and as f, take the distance from the a particle u to some fixed point in the Universe, then we can conclude that the distances between particles in the Universe are bounded by a constant C. So, the Universe is *finite*. Similarly, if we take a time interval between the events as f, we can conclude that the Universe has a *finite lifetime*.

Why particles with large masses do not exist. If we take mass of the particle as f, then we can conclude that the masses of all particles are bounded by some constant C. This result explains the following problem:

- Several existing particle classification schemes allow particles with arbitrarily large masses [3]. E.g., in Regge trajectory scheme, particles form families with masses $m_n = m_0 + n \cdot d$ for some constants m_0 and d. When $n \to \infty$, we have $m_n \to \infty$.
- Only particles with relatively small masses have been experimentally observed (see, e.g., [15]).

These particles with large masses, that are difficult to wed out using equations only, can be easily weeded out if use the notion of "not abnormal".

Dimensionless constants are usually small. This is the reason why engineers and physicists can safely estimate and neglect, e.g., quadratic (or, in general, higher order terms) in asymptotic expansions, even though no accurate estimates on the coefficients on these terms is known [6]. In particular, such methods are used in quantum field theory, where we add up several first Feynman diagrams [4]; in celestial mechanics [16], etc.

7 Application: Chaos Naturally Appears

Restriction to not abnormal also explains the origin of chaotic behavior of physical systems. In mathematical terms, chaos means, in particular, that after some time, the states of the system get close to the so-called *strange attractor*, i.e., a set whose sections are *completely disconnected set.*

Definition 4 *A set S in a metric space X is called* completely disconnected *if for every $s_1, s_2 \in S$, if $s_1 \neq s_2$, then there exist open sets S_1 and S_2 such that $s_1 \in S_1$, $s_2 \in S_2$, $S_1 \cap S_2 = \emptyset$, and $S \subseteq S_1 \cup S_2$.*

In other words, every two points belong to different topological components of the set S. The relationship between this definition and typical elements is given by the following result:

Proposition 4 *In a definable separable metric space, the set of typical elements is completely disconnected.*

Proof Let $s_1, s_2 \in T$, and $s_1 \neq s_2$. Since both points s_i are definable, the distance $d \stackrel{\text{def}}{=} d(s_1, s_2)$ between them is definable as well. Let us take $A_n \stackrel{\text{def}}{=} \{x \,|\, d/2 < d(s_1, x) \leq d/2 + 2^{-n}\}$; then $\cap A_n = \emptyset$, hence there exists an N for which $A_N \cap T = \emptyset$, i.e., for which for every $x \in T$, either $d(s_1, x) \leq d/2$ (hence $d(s_1, x) < d/2 + 2^{-(N+1)}$) or $d(s_1, x) > d/2 + 2^{-N}$. Thus, the desired property holds for $S_1 = \{x \,|\, d(s_1, x) < d/2 + 2^{-(N+1)}\}$ and $S_2 = \{x \,|\, d(s_1, x) > d/2 + 2^{-N}\}$: clearly, $s_1 \in S_1$, $s_2 \in S_2$, $S_1 \cap S_2 = \emptyset$, and $S \subseteq S_1 \cup S_2$.

So, if we assume (as physicists do) that abnormal states are impossible, then we immediate arrive at the chaotic dynamics.

8 Application: Justification of Physical Induction

From the viewpoint of an experimenter, a physical theory can be viewed as a statement about the results of physical experiments. If we had an infinite sequence of experimental results $r_1, \ldots, r_n, \ldots$, then we will be able to tell whether the theory is correct or not. So, a theory can be defined as a set of sequences $r_1, r_2, \ldots$ that are consistent with its equations, inequalities, etc. In real life, we only have finitely many results $r_1, \ldots, r_n$, so, we can only tell whether the theory is *consistent* with these results or not, i.e., whether there is an infinite sequence $r_1, r_2, \ldots$ that starts with the given results that satisfies the theory.

It is natural to require that the theory be *physically meaningful* in the following sense: if all experiments confirm the theory, then this theory should be correct. An example of a theory that is not physically meaningful is easy to give: assume that a theory describes the results of tossing a coin, and it predicts that at least once, there should be a tail. In other words, this theory consists of all sequences that contain at least one tail. Let us assume that actually, the coin is so biased that we always have heads. Then, this infinite sequence does not satisfy the given theory. However, for every n, the sequence of the first n results (i.e., the sequence of n heads) is perfectly consistent with the theory, because we can add a tail to it and get an infinite sequence that belongs to the set $\mathcal{T}$. Let us describe this idea in formal terms.

Definition 5 *Let a definable set R be given. Its elements will be called* possible results of experiments. *By S, we will denote the set of all possible sequences $r_1, r_n, \ldots$, where $r_i \in R$. By a* theory, *we mean a definable subset $\mathcal{T}$ of the set of all infinite sequences S. If $r \in \mathcal{T}$, we say that a sequence r* satisfies *the theory $\mathcal{T}$, or, that for this sequence r,* the theory $\mathcal{T}$ is correct.

Definition 6 *We say that a finite sequence $(r_1, \ldots, r_n)$ is* consistent *with the theory $\mathcal{T}$ if there exists an infinite sequence $r \in \mathcal{T}$ that starts with $r_1, \ldots, r_n$ and that satisfies the theory. In this case, we will also say that the* first n experiments confirm the theory.

Definition 7 *We say that a theory $\mathcal{T}$ is* physically meaningful *if the following is true:*

> *Let r be a sequence $r \in S$ such that for every n, the results of first n experiments from r conform the theory $\mathcal{T}$. Then, the theory $\mathcal{T}$ is correct for r.*

In this case, the universal set consists of all possible infinite sequence of experimental results, i.e., $U = S$.

Proposition 5 *For every physically meaningful theory $\mathcal{T}$, there exists an integer N such that normally, if the first N experiment confirm the theory $\mathcal{T}$, then this theory $\mathcal{T}$ is correct.*

In other words, if the sequence of results r is not abnormal, and the results of first N experiments are consistent with the theory, then the theory is correct. This result shows that we can *confirm* the theory based on finitely many observations.

The derivation of a general theory from finitely many experiments is called *physical induction* (as opposed to *mathematical induction*). These have been many formalizations of different ideas that physicists use, and these formalizations has lead to successful programs that can find a general dependency from the cases. However, in spite of the success in describing several underlying ideas, the general physical induction is difficult to formalize, to the extent that a prominent philosopher C. D. Broad has called the unsolved problems concerning induction a *scandal of philosophy* [2]. We can say that the notion of "not abnormal" justifies physical induction (and thus resolves the corresponding scandal).

Acknowledgments. This work was supported in part by the NASA grant NCC5-209, by the Air Force Office of Scientific Research grant F49620-00-1-0365, by NSF grants EAR-0112968, EAR-0225670, and EIA-0321328, and by the Army Research Laboratories grant DATM-05-02-C-0046.

References

[1] L. Boltzmann, "Bemrkungen über einige Probleme der mechanischen Wärmtheorie", *Wiener Ber. II*, vol. 75, pp. 62–100, 1877.

[2] C.D. Broad, *Ethics and th history of philosophy*, Routledge and Kegan Paul, London, 1952.

[3] L. Brink, M. Henneaux, *Principles of String Theory*, Plenum Press, 1988.

[4] V.B. Berestetsky, E.M. Lifshits, *Relativistic quantum theory*, Pergamon Press, 1974.

[5] R.P. Feynman, *Statistical Mechanics*, W.A. Benjamin, 1972.

[6] R.P. Feynman, R.B. Leighton, M. Sands, *The Feynman Lectures*, Addison-Wesley, 1965.

[7] A.M. Finkelstein, V. Kreinovich. "Impossibility of hardly possible events: physical consequences," *Abstr. 8th Int'l Congr. Log., Methodology & Philosophy of Science*, Moscow, vol. 5, pt. 2, pp. 23–25, 1987.

[8] A.N. Kolmogorov, "Automata and life", In: *Cybernetics expected and Cybernetics unexpected*, Nauka publ., Moscow, 1968, p. 24 (Russian).

[9] V. Kreinovich, I.A. Kunin, "Kolmogorov Complexity and Chaotic Phenomena", *International Journal of Engineering Science*, vol. 41, no. 3–5, pp. 483–493, 2003.

[10] V. Kreinovich, I.A. Kunin, "Kolmogorov Complexity: How a Paradigm Motivated by Foundations of Physics Can Be Applied in Robust Control", In: A.L. Fradkov and A.N. Churilov (eds.), *Proc. Int'l Conf. "Physics and Control" PhysCon'2003*, St. Petersburg, Russia, August 20–22, 2003, pp. 88–93.

[11] V. Kreinovich, L. Longprè, M. Koshelev, "Kolmogorov complexity, statistical regularization of inverse problems, and Birkhoff's formalization of beauty", In: A. Mohamad-Djafari (ed.), *Bayesian Inference for Inverse Problems*, Proc. SPIE, vol. 3459, San Diego, CA, pp. 159–170, 1998.

[12] H.E. Kyburg, Jr., *Probability and the logic of rational belief*, Wesleyan Univ. Press, 1961.

[13] M. Li, P.M.B. Vitanyi, *An Introduction to Kolmogorov Complexity*, Springer, N.Y., 1997.

[14] C.W. Misner, K.S. Thorne, J.A. Wheeler, *Gravitation*, W.H. Freeman, 1973.

[15] Particle Data Group, *Phys. Lett.*, vol. 170B, no. 1, 1986.

[16] L.G. Taff, *Celestial mechanics: a computational guide for the practitioner*, Wiley, 1985.