Coincidences are Not Accidental: a Theorem

Vladik Kreinovich

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

Abstract

In this paper, we formalize and prove the statement that coincidences cannot be accidental, a statement that underlies many useful heuristics in mathematics and physics.

Our proof uses a version of Kolmogorov complexity, a technique originally developed to describe randomness and "accidentalness".

1 Coincidences are not Accidental: A Useful Empirical Fact

Coincidences happen. In mathematics and physics, discoveries often start when someone notices a strange coincidence. For example:

- a large number that has some meaning in one theory is surprisingly equal to the value that was obtained in another theory from quite different reasons, or

- the two functions, that have been defined in two different ways in two different theories, surprisingly coincide, or

- classifications of two different kinds of objects turn out to be surprisingly similar.

Coincidences are useful. Many mathematicians and physicists believe that a coincidence cannot be accidental. So, a theory is created to explain this "empirical" coincidence.

- For example, a crucial point in the development of nineteenth-century analysis was when Gauss observed that the value of a certain elliptic integral coincides with a number defined in a seemingly different way. This coincidence lead him to a general theorem that explained this coincidence and laid the foundation for an essential part of modern mathematics (for a more detailed description of this situation, see, e.g., [1]).

- Similarly, electromagnetic waves were discovered when it turned out that a certain combination of parameters of Maxwell's equations coincides with the velocity of light.

This kind of reasoning is still very fruitful nowadays. To our viewpoint, the most interesting current example of such a fruitful coincidence is presented in the introduction to [5] (see also [4, 10, 2]). The prime divisors of the largest nonstandard simple group (called the "Monster") turned out to coincide with the primes, for which some specific surface has genus 0; the minimal dimension (196,883) in which this group can be realized as a natural transformation group differs only by 1 from the smallest nontrivial coefficient of a seemingly unrelated special mathematical function (called the normalized modular invariant function), etc. All these coincidences, that may seem purely accidental at first glance, lead to a deep mathematical theory that combines:

- finite groups,

- Lie algebras, and

- *conformal quantum field theory.*

The first 40 pages of the monograph [4] are literally filled with coincidences, so that by the moment when the reader reaches page xl, she will agree with the authors' phrase that *"coincidences should not be taken lightly"*.

Coincidences-based heuristics work, but why? For many mathematicians and physicists, there is no doubt that sufficiently nontrivial coincidences should not be accidental. No doubt — but no general theorem as well.

The main goal of the present paper is to present a *theorem* explicitly saying that coincidences are not accidental.

2 Towards a Mathematical Description of the Problem

2.1 What is a coincidence?

The main idea is this: coincidence means that two different definitions lead to the same object (number, function, classification, etc.).

To get more precise, let us remark that any definition of an object can be artificially made more complicated and still define the same object. For example, the number π can be defined as the smallest positive real number x for which $\sin(x) = 0$. We can artificially complicate this definition, e.g., by taking an arbitrary true statement S (e.g., a formula stating that $\sin(y + z) = \sin(y) \cdot \cos(z) + \cos(y) \cdot \sin(z)$ is true for all y and z), and define x as the smallest positive real number for which $\sin(x) = 0$ *and* the statement S is true.

In view of this possibility, it is not really surprising when the same object x has two definitions of *different* length, one shorter and one longer. A true coincidence that deserves our attention is, therefore, when we have two (different) definitions of *the same* length that define the same object.

2.2 What does "not accidental" mean?

When we say that the coincidence between two differently defined objects is not accidental, we mean that there must exist a deeper explanation for this coincidence. This "deeper explanation" may mean two things:

- It may mean that within this same theory, there exists a *third*, simpler (= shorter) *definition* of this same object x (thus explaining the two other definitions as, kind of, complications of this shorter, more basic new definition),

- It may also mean that there exists an *alternative theory* that is more fundamental than the original theory, more fundamental in the sense that:

- each object that is definable in the old theory is also definable in the new one;

- for each definition of an arbitrary object y in the old theory, there is a definition of this same object in the new theory that is *not more complicated* (= not longer) than the given definition in the old theory; and

- at least one object has a *simpler* (= shorter) definition in the new theory.

2.3 To formalize the above idea, we must describe what "definable" means

What "definable" means: in short. The above description uses the word "definable object", "length of the definition", etc. Usually, we say that a formula $F(x)$ with one free variable *defines* an object x if x is the only object that makes this formula true, i.e., if:

- $F(x)$ is true for this object x, and

- $F(y)$ is false for every other object $y \neq x$.

Auxiliary notions that we need to formalize to define what "definable" means. Thus, to formally define this notion, we must describe:

- what an *object* is,

- what a *formula* with one free variable is, and what a *statement* (without free variables) is;

- what is the result of applying a formula $F(.)$ with one free variable to an object x; and

- when is a statement (formula without free variables) *true*.

Preliminary remark: we must fix the alphabet. Objects, formulas, and statements are all representable by sequences of symbols ("words") in a certain alphabet. So, we will assume that some finite alphabet is fixed, and objects, formulas, statements, etc., are words in this alphabet.

Describing the set of true statements. Let us first describe the set of true statements.

In mathematics, to describe which statements are true and which are not, we usually describe two things:

- we describe *axioms*, i.e., statements that we initially assume to be true, and

- we describe *deduction rules*, i.e., rules that enable us to go from some statements, which are already known to be true, to new true statements. For example, if the alphabet contains "and" & with its usual meaning, then from the truth of A and B we can deduce that the statement $A\&B$ is also true.

In mathematical logic, this combination of axioms and deduction rules is usually called a *theory*. When a theory is given, this means that we can eventually generate all possible true statements:

- we start with the axioms;

- we apply all applicable deduction rules to the axioms, and thus get new statements that are true;

- we apply all applicable deduction rules to thus enlarged set of true statements, and, hopefully, conclude that some other statements are true;

- etc.

Thus, the set of all true statements possesses an algorithm that generates, one by one, all the elements of this set (this set may be infinite, so it may take infinitely long to generate all the elements from this set). In mathematical logic and theory of computing, such sets are called *recursively enumerable* (*r.e.*, for short).

There exist more complicated methods of describing a theory, but in all these methods, the set of true statements is r.e.

So, by a set of true statements, we will mean a r.e. set of words.

What are formulas and statements? Some theories have an additional algorithmic property that they are *decidable*, i.e., that there exists an algorithm that, given a statement, checks whether this statements is true or not. Most non-trivial theories are *not* decidable. But even in these theories, while it is algorithmically *impossible* to check whether a statement is true or not, it is usually algorithmically *possible* (and easy) to determine whether a given combination of symbols is indeed a correctly build statement.

E.g., in the first order theory of real numbers, one can easily check that $\exists x(x \cdot x = x)$ is a grammatically correct statement, while $\exists x(x =$ is not.

Thus, by a set of statements, we will mean a decidable subset of the set of all words. Similarly, the set of all formulas with one free variable is another decidable subset of the same set of all words (that does not intersect with the first one).

Only statements (i.e., formulas without free variables) can be true; a formula $F(.)$ with a free variable can be true or false depending on what we apply it to.

What are objects. Objects are usually defined as constructions that satisfy certain properties. Usually, we have an algorithm that generates all possible objects; e.g., the set of all natural numbers can be defined if we start with 0 and 1 and add 0's and 1s. Thus, we will define the set of all objects as a r.e. set of words.

In some cases, this set is decidable, but we do not want to lose generality, so, we will not impose this additional requirement.

Substitution. If we have a formula $F(.)$ with one free variable, and an object x, then we can *substitute* the object x into the formula $F(.)$ and get a statement $F(x)$ (i.e., a formula with no free variables). Substitution is an algorithmic (and easy) operation, so, we will assume that there exists an algorithm that transforms a pair, consisting of a formula with one free variable and an object, into a statement (i.e., into a formula without free variables).

Now, we are ready for the formal definitions.

3   Definitions and the Main Result

Definition 1. *Let a finite set $\mathcal{A}$ be given. This set will be called an alphabet, and its elements will be called symbols.*

- *An arbitrary finite sequence of symbols will be called a word. The set of all words will be denoted by W.*

- *For each word $w \in W$, we will denote its number of symbols (length) by $\mathrm{len}(w)$.*

- *By a language, we mean a tuple $L = \langle \mathcal{A}, \mathcal{F}, \mathcal{S}, \mathcal{O}, \mathrm{sub} \rangle$, where:*

  - *$\mathcal{F}$ is a decidable subset of the set W; its elements will be called formulas with one free variable, or simply formulas;*

  - *$\mathcal{S}$ is a decidable subset of the set W for which $\mathcal{F} \cap \mathcal{S} = \emptyset$; elements of the set $\mathcal{S}$ will be called formulas without free variables, or statements;*

  - *$\mathcal{O}$ is a r.e. subset of the set W; its elements will be called objects; and*

  - *$\mathrm{sub} : \mathcal{F} \times \mathcal{O} \to \mathcal{S}$ is an algorithmic (computable) function that transforms each pair (F, x), where $F \in \mathcal{F}$ and $x \in \mathcal{O}$, into a statement $\mathrm{sub}(F, x) \in \mathcal{S}$; for simplicity, we will also denote $\mathrm{sub}(F, x)$ by $F(x)$.*

Definition 2. *Let a language $L = \langle \mathcal{A}, \mathcal{F}, \mathcal{S}, \mathcal{O}, \mathrm{sub} \rangle$ be given.*

- *By a theory, we mean a r.e. subset $T \subseteq \mathcal{S}$.*

- *We say that a statement $S \in \mathcal{S}$ is true in a theory T if $S \in T$.*

- *We will say that a formula $F \in \mathcal{F}$ is true for an object x if the statement $F(x)$ is true (i.e., if $F(x) \in T$).*

Definition 3. *Let a language* $L = \langle \mathcal{A}, \mathcal{F}, \mathcal{S}, \mathcal{O}, \mathrm{sub} \rangle$ *be given, and let* T *be a theory.*

- *We say that a formula* F *with one free variable* defines *an object* x *if* x *is the only object for which this formula is true, i.e., if*

  - $F(x)$ *is true for this* x, *and*
  - $F(y)$ *is not true for any other object* $y \neq x$.

- *If a formula* $F \in \mathcal{F}$ *defines an object* x, *we say that the formula* F *is a* definition *of the object* x.

- *We say that an object is* definable *in a theory* T *if it has a definition in this theory.*

- *We say that a theory* T *is* non-trivial *if infinitely many objects are definable in this theory.*

- *We say that a definition* F *is* simpler *than the definition* F' *if* $\mathrm{len}(F) < \mathrm{len}(F')$.

- *We say that a definition* F *is* not more complicated *than the definition* F' *if* $\mathrm{len}(F) \leq \mathrm{len}(F')$.

Definition 4. *Let a language* L *be fixed. We say that a theory* T' *is* more fundamental *than the theory* T *if the following three conditions are satisfied:*

- *each object that is definable in the theory* T *is also definable in the theory* T';

- *for each definition* F *of an arbitrary object* y *in the theory* T, *there is a definition* F' *of this same object in the theory* T' *that is not more complicated than* F;

- *that there exists an object* y *and its definition in* T' *that is easier than any of definitions of this object* y *in the theory* T.

Theorem. *Let $L = \langle \mathcal{A}, \mathcal{F}, \mathcal{S}, \mathcal{O}, \mathrm{sub} \rangle$ be a language, and let T be a non-trivial theory in this language L. If in this theory, an object x has two different definitions $F \neq F'$ of the same length ($\mathrm{len}(F) = \mathrm{len}(F')$), then:*

- *either in this same theory T, there exists a third definition F'' of the object x that is simpler that the given two,*

- *or there exists a theory T' that is more fundamental than T.*

Comment. This is exactly what we wanted to prove.

The actual proof (as well as its relation with Kolmogorov complexity) is presented in the Appendix to this paper.

References

[1] J. Borwein and P. Borwein, *A dictionary of real numbers*, Wadsworth & Brooks/Cole, Pacific Groves, CA, 1990.

[2] A. J. Feingold, I. B. Frenkel, and J. F. X. Ries, *Spinor construction of vertex operator algebras, triality, and $E_8^{(1)}$*, American Mathematical Society, Providence, R.I., 1991.

[3] D. Fox, M. Schmidt, M. Koshelev, V. Kreinovich, L. Longpré, and J. Kuhn, "We must choose the simplest physical theory: Levin-Li-Vitányi theorem and its potential physical applications", In: G. Erickson (ed.), *Maximum Entropy and Bayesian Methods*, Kluwer, Dordrecht, 1998 (to appear).

[4] I. B. Frenkel, "Beyond affine Lie algebras", *Proc. Berkeley Intern. Mathem. Congress*, 1986, pp. 821–839.

[5] I. Frenkel, J. Lepovsky, and A. Meurman, *Vertex operator algebras and the Monster*, Academic Press, Boston, 1988.

[6] V. Kreinovich and L. Longpré, "Unreasonable effectiveness of symmetry in physics", *International Journal of Theoretical Physics*, 1996, Vol. 35, No. 7, pp. 1549–1555.

[7] V. Kreinovich and L. Longpré, "Pure Quantum States Are Fundamental, Mixtures (Composite States) Are Mathematical Constructions: An Argument Using Algorithmic Information Theory", *International Journal on Theoretical Physics*, 1997, Vol. 36, No. 1, pp. 167–176.

[8] V. Kreinovich and L. Longpré, "Nonstandard (Non-σ-Additive) Probabilities in Algebraic Quantum Field Theory", *International Journal of Theoretical Physics*, 1997, Vol. 36, No. 7, pp. 1601–1615.

[9] M. Li and P. Vitányi, *An introduction to Kolmogorov complexity and its applications*, Springer-Verlag, N.Y., 1997.

[10] A. Ogg, "A review of [5]", *Bulletin of the American Math. Society*, 1991, Vol. 25, No. 2, pp. 425–432.

Appendix: Proof and its Relation to Kolmogorov Complexity

A1. Proof

In order to prove the theorem, let us first re-formulate it. What out theorem basically says is that if in some non-trivial theory T, an object x has two different definitions $F \neq F'$ of equal length and no simpler (=shorter) definitions, then there exists another theory T' that is more fundamental than T.

Auxiliary result. To construct the desired new theory T', let us first show that in T, there exists an object whose shortest possible definition is longer than $\text{len}(F)$.

Indeed:

- On one hand, there are finitely many words of length $\leq \text{len}(F)$, and thus, only *finitely many* different objects can be defined by such "short" formulas.

- On the other hand, the theory T is assumed to be non-trivial, in the sense that *infinitely many* objects are definable in this theory.

Therefore, there exist an object z whose every definition in T is longer than F.

Construction of a new theory. Since the formula F' defines x, we conclude that $F'(x) \in T$ and $F'(z) \notin T$. Let us now define T' by replacing $F'(x)$ with $F'(z)$ in the theory T.

It is easy to check that T' is a r.e. set, i.e., a theory.

Three conditions that we need to prove in order to show that the new theory is more fundamental. Let us show that T' is indeed more fundamental than T. According to the definition of a more fundamental theory, we must prove three things:

- that each object that is definable in the theory T is also definable in the theory T';

- that for each definition F of an arbitrary object y in the theory T, there is a definition F' of this same object in the theory T' that is not more complicated than F; and

- that there exists an object y and its definition in T' that is easier than any of definitions of this object y in the theory T.

Proving the first condition. Let us first prove the first condition, that each object that is definable in the theory T is also definable in the theory T'. Indeed:

- The object x is definable in T', namely, it is definable by the property F. Indeed:

  - $F(x)$ is still true (i.e., $\in T$), and
  - for every $y \neq x$, $F(y)$ is still not true.

- The object z is definable by the formula F'. Indeed, since the formula F' was defining x in the original theory T, the only statement of the type $F'(y)$ in the set T was the term $F'(x)$. When we constructed T', we replaced this statement with $F'(z)$. Thus, the only statement of the type $F'(y)$, which is now (in T') true, is $F'(z)$. Hence, in the new theory, the formula F' defines the object z.

- Every other object $y \neq x$, $y \neq z$, which was defined in T by some formula F'', is defined in T' by this same formula.

Proving the second condition. Let us now show that for each definition F of an arbitrary object y in the theory T, there exists a definition F' of this same object in the theory T' that is not more complicated than F:

- Let us first prove this fact for the object x. We have assumed that in T, formulas F and F' are the shortest definition of x (of equal length). Since the formula F still defines x in the new theory T', this definition is, thus, not longer than any definition of x in the original theory T.

- For z, the new theory has a definition F' of length $\text{len}(F)$, while any definition in the old theory T was longer. Thus, for old definition, there is indeed a new simpler (= shorter) one.

- For every other object $y \neq x$, $y \neq z$, every formula F that defines y in the theory T defines the same object in the new theory and therefore, the new definition is not longer than the old one.

Proving the third condition. While proving the second condition, we also proved that for the object z, the formula F' that defines this object T' is easier than any of the formulas that define this same object in the original theory T.

The theorem is proven. Three conditions are proven. Thus, the theory T' is indeed more fundamental than T. The theorem is proven.

A2. Relation to Kolmogorov complexity

Similarity. In our proof, for an object x, we considered the shortest length $\mathrm{len}(F)$ of a formula that defines this object:

$$d(x) = \min\{\mathrm{len}(F) \,|\, F \text{ defines } x\}.$$

Thus defined notion is very similar to the notion of *Kolmogorov complexity* (see, e. g., [9]): the Kolmogorov complexity $K(x)$ of an object x is defined as the shortest length of a program that computes x. To get our notion from Kolmogorov complexity, we must:

- replace *programs* with *formulas*, and

- replace *computing* with *defining*.

Similarity explained. The similarity between our notion and Kolmogorov complexity is not accidental:

- one of the main objectives of introducing the notion of Kolmogorov complexity was to formalize the notions of *randomness*, "accidentalness" (see, e.g., physical applications in [9, 6, 7, 8, 3], and

- formalizing of "accidentalness" is exactly what we are interested in.

Possible research directions. In view of the meaningfullness of our analogue of Kolmogorov complexity, it may be desirable to analyze this analogue as the Kolmogorov complexity itself has been analyzed [9]. Hopefully, this analysis will lead to new physically meaningful applications.