An algorithmic approach to fundamental physics

Inés Samengo

For many centuries mathematics was believed to formalize the abstract properties of existing entities. Integer numbers characterised collections of discrete objects. Euclidean geometry described the space we live in. Derivatives represented velocities. The systematisation of such properties was built in terms of axioms, the validity of which was self evident, because axioms attested attributes of the real world. To derive theorems, the supposedly unavoidable rules of logic were used. This context remained stable for many centuries with only marginal questionings, until the 19$^{\text{th}}$ century, when fairly independently Gauss, Schweikart, Bolyai, Lowachewsky and Riemman explored the consequences of replacing Euclid's 5$^{\text{th}}$ axiom (the least self-evident of the five) by one of several alternative axioms [1]. Non-Euclidean geometries were thus born, and even though nobody was sure which aspect of reality they represented, if any, it became clear that geometry could be consistent and interesting all the same. Eventually, more sophisticated physical theories made use of non-Euclidean geometries. But in the meanwhile, the requirement on axioms to be self-evident was called into question.

Mathematicians and logicians had so far worked with statements that could be either true or false. The negation of a true statement is false, and the negation of a false statement is true. The truth of a mathematical statement that refers to existing entities had historically been granted from physics, under the assumption that the world we live in makes sense, so false statements cannot have a physical instantiation. Within this notion of truth, however, one must remain neutral about the truth value of statements that neither themselves nor their negations describe physical processes. The exploration of non-Euclidean geometries launched the exploration of formal systems based on axioms whose truth value could not be based on reality. From that moment on, mathematics focused on the derivation of theorems from postulated assumptions, irrespective of whether the postulates had a hardware instantiation in the physical world. The validity of the theorems hinged upon the validity of the axioms, which was no longer a matter of discussion. Any set of axioms was acceptable (though not necessarily interesting), as long as the derived theory was consistent, that is, as long as no two theorems asserting opposite statements could be demonstrated within the theory.

The unrest turned up a notch with the work of Frege [2], Hilbert [3], and Whitehead and Russel [4], who not only encouraged mathematical derivations dissociated from any corporeal instantiation, but also, detached the act of deducing theorems (the bare use of mathematical logic) from any justification based on self evidence. The laws of inference had to be formulated as equations, and any set of equations giving rise to consistent theories was considered valid. A well founded mathematical theory was then equated with a formal system, that is, a set of formation rules, of axioms, and of inference rules expressed in symbolic language. Formation rules define the statements that the theory deals with. They stipulate the symbols and the grammar with which symbols must be combined to construct strings. Any string not complying with the formation rules is not to be analyzed at all, it can be neither true nor false within the theory, it is simply nonsensical. Among the valid strings, those that can be derived from the theory are considered true (under the premise that the axioms are true). If the inference rules contain a negation, then the negation of the derivable strings are the false statements. Inference rules regulate the norms of logical deduction, allowing the enunciation of new well-formed strings as the logical consequence of old well-formed strings. These rules provide a mechanism for deriving theorems. Axioms are the thematic ingredient of the theory. They are well-formed strings, from which all other strings (theorems) are derived. Once the formal system is stipulated, with its formation rules, inference rules and axioms, the process of deriving theorems can be automatised.

The true statements of a theory (the collection of axioms and theorems) conform a subset of all the statements that comply with the formation rules. In 1931, the mathematical community was shocked when Kurt Gödel demonstrated that consistent and complex formal systems (complex enough as to contain arithmetic) contain well-formed statements whose truth value cannot be assessed within the theory [5]. These statements cannot be deduced from the axioms, and neither can their negations. They make meaningful assertions (and so do their negations), but are neither theorems nor negation of theorems. With varying degree of despair or enthusiasm, mathematicians have been forced to stomach (or savour) this logical oddity ever since.

Two formal systems are equivalent if they give rise to the same set of theorems. When more than a formal system is available, the one with fewest axioms and inference rules is preferred, since it embodies the same truths in fewer statements.

In this essay, I assume that the challenge of constructing physical theories compatible with the reality we observe is analogous to developing a formal system that can produce a fixed set of given theorems as output. In mathematics, theorems are the consequences of formation rules, axioms and inference rules. In physics, the experimental evidence takes the role of theorems, and the task is to solve an inverse problem: Which is the most fundamental formal system from which these theorems come out as consequences? We are looking for formation rules, inference rules and axioms from which the quantities that we measure throughout our lives unfold.

As we live, we gather information of the world around us. We are endowed with the faculty of storing our experiences in memory (or in paper, CDs, or similar devices), so we are able to keep a record of what happens. Of crucial importance to the construction of theories is that we notice regularities. Our perceptions are not completely random. If I let go of my pen, it falls to the ground. If I pick it up and let go of it again (surprise, surprise) it follows essentially the same path as before. There seems to be a pattern in the events we perceive. It therefore makes sense to try to represent them by a more compact description than their raw enumeration, that is, by a formulation that captures the regularities. Mathematical formal systems may serve as an inspiration [6]. Let us represent our evolving perceptions as sequences of variables. Each sequence is a string of symbols. The variables that we perceive at a given time constitute a state vector, and by sampling these variables at different times $t_1, t_2, \ldots, t_n$, we construct a sequence of state vectors. For example, we may choose to measure the position and momentum of a collection of objects at different moments. Alternatively, we may choose to measure the colour of one object, and the weight of another. Different choices of variables constitute different possible definitions of state vectors. Some choices will turn out to be more fruitful than others. In fact, part of constructing a theory is finding a convenient definition of the space of state vectors, compact enough to avoid redundancies, and detailed enough to enable predictions. This part is analogous to selecting a set of formation rules.

A second ingredient in the construction of a theory is to come up with a description, or a representation, of the list of perceived state vectors. In order to be able to compare the benefits and drawbacks of alternative descriptions, it is important to consider them all in a unified format. One way to standardise formats is to define descriptions as computer programs that produce the sequence of measured state vectors as outputs. Different descriptions are hence different computer programs. To keep the notation uniform, we consider programs written in the language used by Turing machines [7, 8], but any computer language is valid, since the translation from one language to another involves a fixed set of rules that do not depend on the description at hand.

One possible description is the explicit list of all the measured variables. The program is simply an ordered pile of “print” instructions, one vector at a time. If the sequence of measured vectors contains no regularities at all, then there is no alternative but to quote all the values, to describe the sequence literally. In the presence of regularities, however, there are much more economic strategies to describe the data. A conceptual description is any such strategy. For a fixed sequence of state vectors, the notion of Kolmogorov complexity [9] allows us to postulate that if descriptions A and B generate the same vectors, A is more conceptual than B if its computer program is shorter. The only way a computer program can be shortened and still produce the same output is by capturing some fundamental law that governs the data. The shorter the program, the larger the insight into the underlying rules. In just a few lines, a lot is explained. The regularities that appear over and over again in the data are compressed into a mere rule.

As stated above, we have the freedom to choose the variables that constitute our state vectors, so we can opt for different formation rules as we please. One real-life situation represented by a given set of variables may perhaps be describable with a very short program, whereas the same situation may only be described by an explicit literal-like program when a different set of formation rules is employed. For example, if we are viewing a sphere falling down from the top of the leaning tower of Pisa, we may chose to describe the vertical position and the vertical momentum of the sphere's center of mass at different points in time, measured with respect to the ground. Given the initial position and momentum, a short program based in Newtonian mechanics will produce all future values on the list. Alternatively, we may choose to list the distance of the sphere with respect to a rambling butterfly. In this case, only a long program will work. Choosing the right variables is as essential an ingredient in the search for conceptual programs as programming cleverly.

Evidence has shown that certain formation rules have the remarkable property of enabling the construction of short programs reproducing the data collected in a large variety of situations, and in repeated trials of one specific situation. Moreover, with these endorsed formation rules, often the short programs obtained in different data sets are similar. Classical physics, for example, is a hugely successful tool to construct short programs. No matter whether we are describing celestial bodies, billiard balls, sound waves or electrical charges, classical physics succeeds in reproducing a large amount of measured data. When changing from one situation to another, a certain part of the programs remains unaltered. Only the initial conditions defining the number of particles, a few of their properties (mass, charge, etc.) and their initial momenta and positions vary. The part of the program that calculates the resulting evolution rests intact. Such programs can be dissected into two sections: the section that varies from data set to data set (the specification of the system and its initial conditions), and the section that remains unchanged (the laws of evolution in classical physics). Strictly speaking, the first section need not be *initial* conditions. They may be final conditions, or half initial and half final, or even other choices. But in any case, they must suffice to determine the situation at hand uniquely. For shortage of notation, here I call all such specifications "initial conditions".

If we have managed to find a set of formation rules, and a criterion to construct short programs that are applicable to a large variety of situations, and that can be dissected into initial conditions and evolution laws, then we have formulated a theory. The (inverse) length of the programs determines the conceptual depth of the theory, and the breadth of the situations in which the theory works determines its generality. I will here consider a theory to be fundamental if it is conceptual and general. The most fundamental theory should be the one that produces the shortest programs, and that can be applied to the largest number of situations.

In mathematics, a formal system defined by a set of formation rules, inference rules, and axioms, spits out theorems. In physics, formation rules define the space inside which we measure the data. Inference rules are analogous to evolution laws, and axioms are analogous to initial conditions. In mathematics, if we change the axioms and keep the same formation and inference rules, the theorems change. In physics, if we change the initial conditions, the trajectory of the system in the relevant space changes, and so does the list of numbers produced by the program. The essential difference between the two is that mathematics is developed forward (from formal systems to theorems) and physics backwards (from evidence to theory).

According to this scheme, a fundamental theory must fulfil two requirements: being conceptual and being general. Two requisites may be complied with in different degrees, so in general, it is not possible to rank all theories in increasing fundamentality. We may still rank them into two separate lists, by their conceptual depth, and by their generality. At least, if the sets of data employed to test different theories have some degree of overlap (it is difficult to rank theories that describe phenomena in completely different domains). For example, a theory formulated in terms of the electroweak interaction is more conceptual than another one written in terms of the weak force and the electromagnetic force. When supersymmetry unified the fate of bosons and fermions, the description became more conceptual (shorter program). In turn, special relativity is more general than classical mechanics, because it contains the latter as a limiting case, and can also describe situations in which classical formulations simply do not produce the right output. General relativity is even more general. But nothing guarantees that as we rank theories according to conceptual depth on the one hand, and according to generality on the other, both orderings produce the same hierarchy. Some physicists hope, however, that even if partially successful theories may not be completely ordered, there is one yet-to-be-discovered theory that is the most conceptual and the most fundamental of all, the so-called Theory of Everything (or also, Theory of Nothing [10, 11]). The hope is that this ultimate theory can be applied to absolutely all situations, and is trivially simple to formulate.

In the meanwhile, we must struggle with the messy zoology of theories that fill our libraries, colleges and common knowledge. These theories vary greatly in their domain of applicability and scope. Some of them, typically quite restricted in their ability to reproduce hard quantitative data, aim at predicting qualitative features of large-scale processes, as human behaviour (emotional reactions, decision taking, language production, etc.) or social processes (diasporas, finances, information flow in communication devices, etc.). Other theories, often formulated in precise mathematical terms, deal with atoms, electrons and chemical reactions. To date, no single theory can explain all the data. Even if quantum mechanics provides a seemingly accurate description of fields and particles, any attempt to predict social processes from fundamental physics is deemed to fail, not only because we are still in the dark about many of the links connecting different scales, but also, because even if we knew them all, the chaotic nature of complex systems makes low-level modelling utterly useless. Nevertheless, the scientific community often regards disciplines as hierarchically ordered. Computer programs written in terms of physical laws can describe atoms fairly accurately, and even also some simple molecules, providing the theoretical bases from which the more heuristic laws of chemistry take off. Programs written in terms of chemical principles, in turn, can be stretched up to describe fairly complex molecules, from which the basic laws of biology take over. The building of science can be iteratively assembled, going through biology, medicine, neuroscience, psychology, sociology and so forth. As we zoom out, we must rewrite the programs, in terms of more macroscopic laws. In every transition, however, there is a certain intermediate zone where both macro and micro descriptions coexist, and sometimes there is even a precise theoretical derivation that mediates the handover. Therefore, although physics cannot explain psychological data, there seems to be some optimistic hope that ultimately, the whole universe is ruled by a coherent collection of laws that are applicable universally. Even if due to numerical issues, and also to some yet missing theoretical links, no computer program has managed to prove it.

Before we dive into even more theoretical matters, a note of caution is in order, since practical issues may well bar the evaluation of the success of a theory. So far, we have taken the list of data obtained in experiments as certain. However, measurements typically involve some degree of uncertainty. In these circumstances, there is no clear-cut procedure to determine whether the output of a program coincides with the measured data. This uncertainty becomes unavoidable when the state variables are described by real numbers, most of which cannot be measured with infinite precision, and most of which are incomputable, that is, impossible to represent by a Turing machine. The sequence of measured values, hence, must be interpreted probabilistically, with probabilities understood in their Bayesian conception: as a measure of the certainty with which a given observer trusts the reported data. The comparison between the output of the program and the measured data is therefore also probabilistic. As a consequence, the evaluation of theories is infused with fuzziness. Once the sequences of data collected by experience are not intended to be matched with infinite precision, theories are allowed to provide approximate descriptions of reality. In these circumstances, it is difficult to rank the conceptual depth of two theories if, for example, one of them has a shorter program, but the other matches the measured data more accurately. How should generality be evaluated, if the data generated by the program match the experiments with a precision that varies from one situation to another? The result of any comparison depends on the muddy criteria with which we weigh the trade-off between conceptual depth + generality, against accuracy. Indeed, if we allow ourselves to be lax in the comparison with reality, it is easy to fall onto extremely fundamental theories, grounded on the idea that basically anything can happen. It could be argued that string theory, with its 10^{500} possible landscapes, is essentially one such theory. So little is determined by the corresponding program, that its output consists in a vast collection of possible universes, with only one, or a few, matching the data produced in our actual world. On the other hand, by being excessively strict, we may easily fall on the opposite conundrum, equally inconvenient. If we demand the program to produce data that match our noisy experiments precisely, only literal list-like descriptions pass the test. No theoretical abstraction can be distilled from them. This difficulty permeates the whole construction of theories, but for the moment, I will assume that some criterion has been found, to determine in which situations the output of a program matches the data acceptably.

Humans, and probably many other species, construct and have constructed theories ever since, probably as an economic way to predict what will happen around them, and by adjusting their response with anticipation, boosting their fitness and survival. Fundamental theories, however, provide us with more than an efficient way to survive and pass our genes on (not to mean that these advantages are trifle). They also enable us to identify the basic elements of reality. We tend to believe that the theoretical entities that enable the calculations performed by the program have a fundamental ontological status. For example, classical physics is formulated temporally and spatially: all events have spatiotemporal coordinates. Irrespective of whether such coordinates are provided as output of the program or not, once we become accustomed to a successful theory, it is natural to believe that time and space exist. Simply because by postulating their existence we may reproduce a lot of the observed data. The same holds for mass, energy, momentum, and for all the intermediate entities that we do not necessarily observe, but that are generated by the program as a means to interpolate between the events we do observe. Our natural tendency to think about the world in a compact and conceptual fashion drives us to believe that fundamental physics describes the ultimate reality, and that such reality exists even when we do not observe it.

This is one of the reasons why quantum mechanics provoked such a dramatic upheaval, when it came on stage. The experiments performed at truly microscopic scales produced results that could not be predicted by any of the previous theories. The first reaction was to conjecture that we might still not have the correct theory [12]. The commotion escalated when theoretical and experimental analyses concluded that the unpredictability of measurements was not due to lack of knowledge, but rather constituted an essential ingredient of nature [13]. Moreover, not only future events were revealed to be fundamentally undefined, but also past ones, even if macroscopic and seemingly irreversible [14, 15]. Quantum mechanics formally declares that it is literally impossible to construct a program from which we can predict all what we measure. At least part of what we observe is intrinsically random, and any attempt to describe it requires some degree of literal enumeration. One could of course claim that this is a shortcoming of quantum mechanics, and hope that there might be yet some other formulation that achieves the goal that quantum mechanics fails to deliver. Only thing, no such theory has been found. None of the previous deterministic formulations are able to predict the outcomes of microscopic experiments, and no deterministic theory can comply with Bell's inequality [13]. So we are left with two alternatives: Either we embrace this puzzling quantum mechanics, or we embrace no theory at all, leaving us with nothing but a list of arbitrary numbers to describe reality. And even though quantum mechanics is not able to construct a compact program that generates the data we observe in single trials, it does succeed in describing something even broader: the probability distribution with which results would be obtained, were we to test the same situation repeatedly. It also provides an accurate description of the correlations between different variables (often through entanglement), even if their actual values may fluctuate [16]. Moreover, when dealing with macroscopic objects and partial knowledge (decoherence) it generates the laws of classical mechanics as a limiting case [17]. This is less than what we had began hoping for, but certainly more than nothing.

If we decide to embrace quantum mechanics, we must be ready to accept that the variables enabling conceptual programming are defined in terms of probability amplitudes. When attempting to reproduce data collected in individual trials, long (non-conceptual) programs are required, since the theory cannot predict single samples. Quantum mechanics does not model what happens, but rather, whatever could have happened, including the interferences between the alternative possibilities. In the derived ontology, reality becomes an attribute of collections of entangled histories, and not of individual experiences.

The ontological shift modifies our way to construct theories. We must now relinquish the aspiration to program the outcome of individual experiences, we must content ourselves with programs that reproduce the statistics of the data observed in repeated experiments. This requirement enforces a modification of the formation rules. The experiences we collect throughout our (single) life do not form a valid string, nor a target for a theory. They are only a tiny portion of a much larger string, that contains all what we could have experienced, and the resulting interferences.

In quantum mechanics, when an observer makes a measurement, different versions of him or her become entangled with the different possible outcomes of the experiment [18, 19]. Each of these versions (that is, each branch of the wavefunction) contains one observer reporting one of the many possible experimental results. For each such observer, the density matrix of the system loses its off-diagonal terms, and from that moment on (or back), histories cannot interfere with each other. Interference is still possible, but it cannot be detected by the observers that performed the measurements. The moment the measurement was done, the subset of histories that included interference was eliminated from the menu of histories that are accessible to the measuring agents. Yet, a higher-level observer monitoring the histories of the lower-level ones may have access to interferences. If he or she is careful to look into the system only after the lower-level measurement, no pruning of histories takes place. The whole collection will therefore evolve with a density matrix that still contains the diagonal terms. As a consequence, in quantum mechanics, there is no such thing as one story: Different observers report different histories of the same situation. Instead, in the ontology of classical mechanics, things happen whether we observe them or not. We may choose to collect data about the system any time, and all choices produce results that are compatible with one and the same story.

In this essay, I have proposed a certain parallelism between mathematical and physical theories, equating the construction of theories with the search for formal systems whose formation rules, inference rules and axioms give rise to the world we live in. The goal is not only to predict numbers, but also, to understand the nature of the processes we describe. Such understanding calls for fundamental theories, here defined as those supporting short programs with wide applicability. The parallelism between physics and mathematics may seem somewhat outmoded. For more than a century mathematicians have been endorsing formal systems with no physical counterpart, and here I support a search for a formal system whose instantiation is the world we live in. This search commends the old notion of truth, in which statements are deemed true if they describe entities existing in the real world. I could potentially consider endorsing such a notion of truth to describe processes governed by classical physics, since for them, there is such thing as **one** reality. That one history can be considered the true one, and all other alternative evolutions, not actually happening, are false. This notion of truth, however, cannot be arrogated to all quantum phenomena, because in the quantum realm there is no single reality. The system and the observer become entangled, and what happens to the system depends on what happens to the observer. Measuring therefore includes self-referential components, and there is no absolute way an observer living inside the universe can assign universal truth values to histories, nor to states. This indeterminacy is reminiscent of Gödel's incompleteness theorem. With varying degrees of despair or enthusiasm, we may have to stomach or savour the fact that the reality of histories (including our own) cannot be determined from inside the universe, just as the truth of many mathematical theorems cannot be deduced from inside the formal system they belong to. A striking humility lesson, if there is one.

References

[1] Nagel E, Newman JR. Gödel's Proof. New York: New York University Press; 1958.

[2] Frege G. Begriffsschrift: eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Halle: L. Nebert; 1879.

[3] Hilbert D. Grundlage der Geometrie (3rd edition). Leipzig: Teubner; 1909.

[4] Whitehead AN, Russell B. Principia Mathematica (3rd edition). Cambridge: Cambridge University Press; 1962.

[5] Gödel K. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I. Monatshefte für Mathematik und Physik. 1931;38(1):173198.

[6] Tegmark M. Our Mathematical Universe. New York: Alfred A. Knopf; 2014.

[7] Turing AM. On Computable Numbers, with an Application to the Entscheidungsproblem. Proceedings of the London Mathematical Society. 1936;42(1):230–265.

[8] Turing AM. On Computable Numbers, with an Application to the Entscheidungsproblem: A correction. Proceedings of the London Mathematical Society. 1938;43(6):544–546.

[9] Kolmogorov AN. On Tables of Random Numbers. Sankhya: The Indian Journal of Statistics, Ser A. 1963;25:369–375.

[10] Standish R. Theory of Nothing. Createspace Independent Pub.; 2006.

[11] Gefter A. Trespassing on Einstein's Lawn. New York: Bantam Books; 2014.

[12] Einstein A, Podolsky B, Rosen N. Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? Physical Review. 1935;47(10):777–780.

[13] Bell J. On the Einstein Podolsky Rosen Paradox. Physics. 1964;1(3):195–200.

[14] Wheeler JA. The "Past" and the "Delayed-Choice" Double-Slit Experiment. In: Marlow AR, editor. Mathematical Foundations of Quantum Theory. Academic Press; 1978. p. 9–48.

[15] Kim YH, Yu R, Kulik SP, Shih Y, Scully MO. A Delayed "Choice" Quantum Eraser. Physical Review Letters. 2000;84(1):1–5.

[16] Rovelli C. Relational quantum mechanics. International Journal of Theoretical Physics. 1996;35(8):16371678.

[17] Zurek WH. Decoherence and the transition from quantum to classical. Physics Today. 1991;44(10):36–44.

[18] Wigner EP. Symmetries and Reflections. Woodbridge: Ox Bow Pr; 1967.

[19] Everett H. The Many-Worlds Interpretation of Quantum Mechanics. Princeton University: PhD dissertation; 1957.