1-2020

Need for Simplicity and Everything Is a Matter of Degree: How Zadeh's Philosophy is Related to Kolmogorov Complexity, Quantum Physics, and Deep Learning

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Olga Kosheleva
The University of Texas at El Paso, olgak@utep.edu

Andres Ortiz-Muñoz
California Institute of Technology, aoortiz@caltech.edu

Recommended Citation

Kreinovich, Vladik; Kosheleva, Olga; and Ortiz-Muñoz, Andres, "Need for Simplicity and Everything Is a Matter of Degree: How Zadeh's Philosophy is Related to Kolmogorov Complexity, Quantum Physics, and Deep Learning" (2020). *Departmental Technical Reports (CS)*. 1407.
https://scholarworks.utep.edu/cs_techrep/1407

Need for Simplicity and Everything Is a Matter of Degree: How Zadeh's Philosophy is Related to Kolmogorov Complexity, Quantum Physics, and Deep Learning

Vladik Kreinovich, Olga Kosheleva, and Andres Ortiz-Muñoz

Abstract Many people remember Lofti Zadeh's mantra – that everything is a matter of degree. This was one of the main principles behind fuzzy logic. What is somewhat less remembered is that Zadeh also used another important principle – that there is a need for simplicity. In this paper, we show that together, these two principles can generate the main ideas behind such various subjects as Kolmogorov complexity, quantum physics, and deep learning. We also show that these principles can help provide a better understanding of an important notion of space-time causality.

1 Need for Simplicity and Everything Is a Matter of Degree: Two Main Principles of Zadeh's Philosophy

Two main ideas of Zadeh's philosophy: in brief. In a nutshell, the main principle behind Lotfi Zadeh's research is a phrase that he himself repeated many times: Everything is a matter of degree. However, it is easy to forget – especially for mathematicians by training (like us) – that he also actively pursued another principle: the idea that there is a need for simplicity.

These two principles underlie his ideas of fuzzy logic: yes, the first principle is that everything is a matter of degree, so instead of 0 (false) and 1 (true) we have all possible numbers from the interval $[0, 1]$. However, to flesh out this argument, Zadeh used simplicity as a guiding principle – thus, he selected the simplest "and"- and "or"-operations (i.e., t-norms and t-conorms), he selected the simplest membership functions, etc.; see, e.g., [13].

Vladik Kreinovich and Olga Kosheleva
University of Texas at El Paso, El Paso, Texas 79968, USA
e-mail: vladik@utep.edu, olgak@utep.edu

Andres Ortiz-Muñoz
California Institute of Technology, Pasadena, California 91125, USA
e-mail: aoortiz@caltech.edu

Many of us remember that every time we came to him with some complex developments, his first reaction always was: can we make it simpler?

What we do in this paper. At first glance, it may seem that these two principles – the need for simplicity and everything is a matter of degree – are specific for fuzzy logic and related areas. However, what we show in this paper is that these principles are valid way beyond fuzzy logic – i.e., beyond the description of imprecise natural language statements.

We show how these principles can be naturally applied:

- first, to the principles themselves,
- then to the description of the physical world, and
- finally, to the way we gain knowledge about the world.

As a result, we naturally arrive at the main ideas behind, correspondingly, Kolmogorov complexity, quantum physics, and deep learning.

This shows that there probably is a lot of remaining potential in these principles – which will enable us to solve many useful problems in the future as well.

2 Simplicity and Everything-Is-a-Matter-of-Degree Principles Naturally Lead to Kolmogorov Complexity

Let us apply the principles to themselves. What are the possible applications of the above two principle? The first natural idea is to apply these principles to the principles themselves: namely, to the description of what is simple and what is complex.

Need for degree of complexity. In accordance with the everything-is-a-matter-of-degree principle:

- we should not naively divide the constructions into simple and complex;
- instead, we should assign, to each construction, a degree of complexity.

What is the simplest degree of complexity? The complexity of an object is naturally described by how complex it is to design it. In particular, for objects generated by a computer, a natural measure of complexity is the complexity of a program that generated this object.

What is the simplest possible measure of complexity? The simplest possible idea is just to count the number of symbols in a program: if the program if short, it is probably rather simple, and, vice versa, if the program is long, it is probably complicated. So, a natural measure of a program complexity is its length.

For an object, the natural measure of its complexity is thus the length of the program that generates this object. Of course, we can have different programs for generating the same object:

- we can have reasonably short programs that compute this object directly, and

- we can have over-complicated programs that beginning students often write.

In the US, such over-complicated constructions are known as Rube Goldberg machines – after a cartoonist who excelled in describing unnecessarily complex ways of performing simple tasks.

From this viewpoint, to gauge the complexity of an object, we should take not just *any* program that generates this object, but the *simplest* such program – and we decided to measure the program's complexity by its length, the shortest such program. Thus, we arrive at the following definition.

We naturally arrive at Kolmogorov complexity. As we mentioned, the complexity of an object *x* can be measured by the length of the shortest program that generates this object. This definition is known as *Kolmorogov complexity* (see, e.g., [5]) and is usually denoted by $K(x)$.

Comments. Simultaneously with Kolmogorov, similar ideas were proposed by R. Solomonoff and G. Chaitin [5].

The length of a program depends on the programming language. However, as Kolmogorov has shown, the definitions $K_1(x)$ and $K_2(x)$ of Kolmogorov complexity based on two different languages differ by a constant:

$$|K_1(x) - K_2(x)| \leq C_{12}$$

for all *x*. Thus, in effect, different programming languages lead to a similar notion of Kolmogorov complexity.

Why was Kolmogorov complexity invented – and how is it useful? While we showed that Kolmogorov complexity naturally follows from the two main principles of Zadeh's philosophy, this is not how this notion was originally invented. This notion was invented by Andrei Kolmogorov, a 20 century Russian mathematician who was one of the pioneers of modern (mathematically precise) probability theory. In addition to doing pure mathematics, Kolmogorov spent a lot of time and effort on applications. Because of this, he noticed a discrepancy between the traditional mathematical description of randomness and the physicists; and engineers' intuitive ideas of randomness.

Indeed, from the purely mathematical viewpoint, when you flip a fair coin *n* times, all resulting binary sequences – describing the *n* results – have the exact same probability 2^{-n} and are, in this sense, completely equal. However, from the practitioner's viewpoint:

- some of these sequences are random – in the sense that one can expect to get them when flipping an actual coin, and
- some sequences like 000...0 or 010101...01 are *not* random and thus, not expected.

From the purely mathematical viewpoint, if the same person wins the main multi-million prize of a super-lottery several years in a row, it is possible. However, if this would happen in real life, it would be clear that the lottery is rigged.

To eliminate this discrepancy, Kolmogorov decided to formalize the intuitive notion of randomness of a finite sequence. His idea was very straightforward:

- Why is a sequence 000...0 not possible? Because this sequence is very simple, it can be obtained by running a very simple program.
- Similarly, a sequence 010101...01 can be computed by a very simple program: just continue printing 01 in a loop.
- On the other hand, if we have a sequence of actual coin-flipping results, there will be hardly any dependencies, so the only way to generate this sequence 011... is to write something like `print`(011...). The length of such a program is about the same as the length $n = \text{len}(x)$ of the original binary sequence.

In other words:

- for a truly random sequence x, the length $K(x)$ of the shortest program that generates this sequence is close to $\text{len}(x)$: $K(x) \approx \text{len}(x)$, while
- for a sequence which are not random, such length is much smaller:

$$K(x) \ll \text{len}(x).$$

This led Kolmogorov and other researchers to define a random sequence as a sequence for which $K(x) \geq \text{len}(x) - C$, for some appropriate constant C.

This definition depends on C.

- If we select C to be too small, we may miss a sequence which is actually random and which accidentally has some dependencies.
- On the other hand, if we select C to be too large, we risk naming not-very-random sequences as random.

So, instead of a single definition of what is random and what is not random, we get, in effect, a *degree* of randomness of a sequence – gauged by the smallest value C for which the above inequality holds. This smallest value, as one can easily check, is equal to the difference $C = \text{len}(x) - K(x)$.

- When this difference is small, the sequence is truly random.
- When this difference is large, the sequence is not so random.

Kolmogorov complexity and the related notion of randomness have indeed been very useful, both in physics applications and in the analysis of algorithms; see, e.g., [5]. And if this notion was not invented by Kolmogorov – it was invented in the 1960s, at the same time as fuzzy logic – Zadeh's principles could have helped invent it.

3 Simplicity and Everything-Is-a-Matter-of-Degree Principles Naturally Lead to the Main Ideas Behind Quantum Physics

Let us apply these principles to the description of the physical world. Fuzzy logic originated by the observation that:

- in most cases, we cannot say that someone is absolutely young or absolutely old;
- for most people, "young" is a matter of degree.

If we apply the same everything-is-a-matter-of-degree principle to the physical world, we conclude that for every two possible states:

- we should not expect the system to be definitely in one state or definitely in another state;
- in many cases, we should have a system which is, with some degree, in one state, and with some degree, in another state.

How can we describe this in precise mathematical form? We have two original states s_1 and s_2. Each state can be described by, e.g., the values of the corresponding quantities, i.e., by a tuple $s_i = (s_{i1}, \ldots, s_{in})$ of real numbers. We need a function that transforms the two states into a new combined state.

From the mathematical viewpoint, the simplest functions are linear functions. Thus, the simplest possible way to combine the two states is to have a linear combination $a_1 \cdot s_1 + a_2 \cdot s_2$. This is exactly what in quantum physics (see, e.g., [2, 12]) is called *superposition* of the two states – with the only difference that in quantum physics, the coefficients a_i can be complex numbers.

The existence of such superposition is one of the main ideas behind quantum physics – and the main idea behind *quantum computing* and its successes; see, e.g., [8].

What is the dynamics here? To fully describe a physical phenomenon, it is not sufficient to describe the current state, we need also to describe how this state changes in time – i.e., what is the corresponding dynamics.

We thus need to describe how the state s of a physical system changes with time, i.e., how the rate of change $\dfrac{ds}{dt}$ depends on the state: $\dfrac{ds}{dt} = F(s)$, for some function $F(s)$.

Which function $F(s)$ should we select? For this selection, it is reasonable to use the simplicity principle. As we have mentioned earlier, the simplest functions are linear functions, so it is reasonable to use linear function $F(s)$ – and this is exactly what is happening in quantum physics, where the Schroedigner's equation describes the rate of change in a state as a linear function of this state [2, 12].

4 Simplicity and Everything-Is-a-Matter-of-Degree Principles Naturally Lead to the Main Ideas Behind Deep Learning

How do we learn? Let us now go from the description of a physical world to the description of how we learn things about this world. First, according to the everything-is-a-matter-of-degree principle, at each given moment of time, about each statement about the world, we do not necessarily have a precise opinion. Mostly, we have a degree of confidence about the given statement.

How do these degrees of confidence change? The simplest possible functions, as we have mentioned several times, are linear functions. So, a natural transformation of degrees is linear.

The problem with using only linear transformations is that by combining several linear transformations, we can only have linear functions – and some real-world processes that we want to describe are not linear. So, we need some non-linear transformations as well.

What are the simplest nonlinear transformations? First, in general, the more inputs, the more complex the transformation. From this viewpoint, the easiest are the functions of one variable.

What are the simplest functions of one variable which are not linear? A natural idea is to use functions which are *piece-wise linear*, i.e., for example, consist of two parts on each of which they are linear. The simplest value separating the two parts is 0, so we end up with a function $s(x)$ which is equal to $c_- \cdot x$ for $x \leq 0$ and to $c_+ \cdot x$ for $x \geq 0$, for some appropriate values c_- and c_+.

What are the simplest real numbers? Clearly, 0 and 1, so we end up with a function which is equal to $0 \cdot x = 0$ for $x \leq 0$ and to $1 \cdot x = x$ for $x \geq 0$. This function can be equivalently described as $s(x) = \max(0, x)$. So, we arrive at the following description of the corresponding computational schemes.

We naturally arrive at deep learning. The input to the non-linear transformer comes from a linear combination of inputs. So, we have transformations of the type

$$x_1, \ldots, x_n \to \max\left(0, \sum_{i=1}^{n} w_i \cdot x_i - w_0\right)$$

for appropriate wights w_i. This is exactly the transformation performed by each neuron in a deep neural network; see, e.g., [3]. Thus, by combining such transformations, we get the main structure of deep neural networks.

If instead of nonlinear functions of one variable, we allow nonlinear functions of several variables, then the simplest are – like in fuzzy logic – min and max. Such transformations form the basis of convolutional neural networks – another component of (and type of) deep learning [3].

So, Zadeh's ideas are in good accordance with deep learning as well.

5 What Next?

What next? There are two natural ideas.

First is to continue applying Zadeh's ideas to other areas – and we will provide an example of such application in the next section.

Another idea is to take into account that a degree does not have to be a single number. We have already seen this in quantum physics – where the coefficients a_i are, in, general complex numbers $a + b \cdot i$, i.e., in effect, pairs of real numbers. Similar ideas of using interval-valued and more general fuzzy degrees – i.e., degrees

described by several numbers – have been successfully in fuzzy applications as well; see, e.g., [6]. And even in applications like education, where traditionally the degree of student's knowledge was described by a single number, the new tendency is to use several numbers, e.g., describing the degree of student's knowledge of different topic – and thus providing a more adequate description of the student's knowledge; see, e.g., [1].

6 Causality as a Matter of Degree

6.1 Motivations

Space-time causality is one of the fundamental notions of modern physics; however, it is difficult to define in observational physical terms. Intuitively, the fact that a space-time event $e = (t, x)$ can causally influence an event $e' = (t', x')$ means that what we do in the vicinity of e changes what we observe at e'. If we had two copies of the Universe, we could perform some action at e in one copy but not in another copy; if we then observe the difference at e', this would be an indication of causality. However, we only observe one Universe, in which we either perform the action or we do not. At first glance, it may seem that in this case, there is no meaningful way to provide an operational definition of causality. In this section, we show that such a definition is possible if we use the notions of algorithmic randomness and Kolmogorov complexity. The resulting definition leads to a conclusion that space-time causality is a matter of degree.

Comment. The main idea of this section first appeared in [4].

6.2 Defining Causality Is Important

Space-time causality is important. Causal relation between space-time events (i.e., points in space-time) is one of the fundamental notions of physics; see, e.g., [2, 7, 12]). Because of this, many fundamental physical theories describe, among other things, the causal relation between space-time events.

According to modern physics, space-time causal relation is non-trivial. In Newton's physics, it was assumed that influences can propagate with an arbitrary speed, constituting, in effect, immediate action-at-a-distance. Under this assumption, an event $e = (t, x)$ occurring at moment t at location x can influence an event $e' = (t', x')$ occurring at moment t' at location x' if and only if the second event occurs later than the first one, i.e., if and only if $t < t'$.

In special relativity, the speeds of all the processes are limited by the speed of light c. In this theory, an event $e = (t, x)$ can influence an event $e' = (t', x')$ if during the time $t' - t$, the faster possible process – light – can cover the distance $d(x, x')$

between locations x and x', i.e., if

$$c \cdot (t' - t) \geq d(x, x').$$

In the general relativity theory, the space-time is curved, so the corresponding causal relation is even more complex. This relation is also complex in alternative gravitation theories; see, e.g., [7].

Need for experimental verification of space-time casuality. Different theories, in general, make different predictions about the causality. So, to experimentally verify fundamental physical theories, we need to be able to experimentally verify the corresponding space-time causality. In other words, we must be able to experimentally check, for every two space-time events a and b, whether the event a can causally influence the event b.

Need for a theory-free verification of space-time causality. Since the space-time causality is fundamental, more fundamental than specific partial differential equations that describe the physical fields and/or their relation with space-time, it is desirable to be able to experimentally check this causality in a theory-free way, without invoking other fields and corresponding differential equations.

In this section, we describe a possible way of such theory-free experimental validation of space-time causality.

6.3 Defining Causality: Challenge

Intuitive meaning of space-time casuality. Intuitively, the fact that a space-time event e can causally influence an event e' means that:

- what we do in the vicinity of e
- changes what we observe at e'.

How to transform this meaning into a definition: a hypothetical idea. The above intuitive meaning of space-time causality can easily lead to a observational definition *if* we had two (or more) copies of the Universe. In this case, to check that e can causally influence e', we could do the following (see, e.g., [9]):

- in one copy of the Universe, we perform some action at e, and
- we do not perform this action in the second copy of the Universe.

If the resulting states at e' are different in the two copies of the Universe, this would be an indication of causal relation between e and e'.

Comment. This interpretation of causality is known as a *counterfactual* interpretation; see, e.g., [10]. This name comes from the usual interpretation of *counterfactual* statements, i.e., statements of the type "If we were born in Sahara, we would have been better adjusted for the warm climate." These statements are called counterfactual because the premise (we are born in Sahara) contradicts to the facts. The usual

interpretation of such statements is to consider not just our world, but also the whole set of possible worlds. To check whether a counterfactual statement is true we select, among all possible worlds in which the premise is satisfied, the one which is the closest to our own world. The statement is considered true if the conclusion holds in this selected world.

Similarly, in the counterfactual interpretation of causality, instead of considering only one world, we consider all possible worlds. We then say that e casually influences e' if in every world in which e occurs, this occurrence affects e'. For example, we want to check whether a rain dance (e) causes rain (e'). In our world, we observe a rain dance, and we observe rain, but we cannot tell whether the rain was caused by the rain dance or not. Intuitively, the way to check is to see if rain dances leads to rain. So, in one possible world, we perform a rain dance, in another possible world, we do not perform it. If, as a result, we see rain in the first world but not in the second one, this is good indication that the rain dance indeed causes rain.

World 1   World 2

$$\text{rain} * e' \qquad e' * \text{no rain}$$

$$\text{rain dance} * e \qquad e * \text{no rain dance}$$
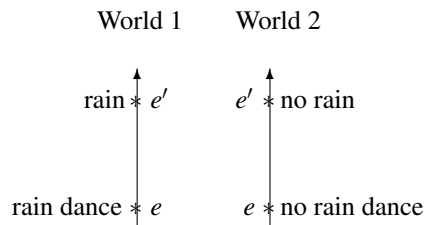
Can we make this idea practical? In reality, we only observe one Universe, in which we either perform the action or we do not.

At first glance, it may seem that in this case, there is no meaningful way to provide an operational definition of space-time causality.

Our idea. In this section, we show that a meaningful operational definition of space-time causality is possible if we use the notions of algorithmic randomness and Kolmogorov complexity. Before we explain our idea, let us briefly recall the corresponding notions.

6.4 Algorithmic Randomness and Kolmogorov Complexity: A Brief Reminder

The corresponding notion of independence. In probability theory, in addition to analyzing what is *random* and what is not, it is also important to decide when the two events are *independent* and when they are not. Once we have two finite binary sequences x and y, the idea that y is independent on x can be described in a similar way:

- if y is independent on x, then knowing x does not help us generate y;
- in contrast, if y depends on x, then knowing x can help us compute y.

For example, if we know the locations and velocities x of a mechanical system at some moment of time t, we can use this information to easily compute the locations and velocities y at the next moment of time $t + \Delta t$. In contrast, an irrelevant information x (e.g., locations and velocities of particles on another planet) does not help in computing y.

To formalize this intuition, we should consider programs that use x as an input to generate y.

Definition 1. *Let a programming language be fixed. By a* relative Kolmogorov complexity $K(y|x)$ *of a finite binary string y in relation to a binary string x, we mean the shortest length of a program that, when using x as an input, generates y:*

$$K(y|x) \stackrel{\text{def}}{=} \min\{\text{len}(p) : p(x) \text{ generates } y\}.$$

Comment. Intuitively, if using x helps to compute y, i.e., if $K(y|x) \ll K(y)$, this means that y depends on x. Vice versa, if using x does not help to compute y, i.e., if $K(y|x) \approx K(y)$, this means that x and y are independent. We can describe this in a way similar to the above definition of randomness:

Definition 2. *Let an integer $C > 0$ be fixed.*

- *We say that a string y is* independent *of the string x if*

$$K(y|x) \geq K(y) - C.$$

- *We say that a string y is* dependent *on the string x if*

$$K(y|x) < K(y) - C.$$

6.5 How to Define Space-Time Causality: Analysis of the Problem and the Resulting Definition

First seeming reasonable idea. At first glance, the above notion of dependence can already lead to a natural definition of space-time causality:

- First, we perform some measurements and observations in the vicinity of the event e. Since most nowadays measuring instruments are computer-connected, each such measurement produces a computer-readable output. In the computer, everything is represented as a sequence of 0s and 1s, so the results of all the measurements and observations will also be represented as a sequence x of 0s and 1s.
- We also perform measurements and observations in the vicinity of the event e', and also produce a sequence x' of 0s and 1s.
- If x' depends on x, i.e., if $K(x'|x) \ll K(x')$, then we claim that e can casually influence e'.

Unfortunately, this idea does not always work. Yes, if e can casually influence e', then we indeed expect that knowing what happened at e can help us predict what is happening at e'. However, the inverse is not necessarily true: we may have identical observations $x = x'$ at events e and e' simply because they are both caused by the same event e'' from the joint past of events e and e'.

For example, if two people at different locations are watching the same movie, then their observations are identical, but not because they causally influence each other, but because they are both influenced by a past event e'' (of making this movie).

How to transform the above idea into a working definition. According to modern physics, the Universe is quantum in nature. For many measurements involving microscopic objects, we cannot predict the exact measurement results, we can only predict probabilities of different outcomes. The actual observations are truly random.

Moreover, for each space-time event e, we can always set up such random-producing experiments in the small vicinity of e, and generate a random sequence r_e. For example, we can locally set up a Stern-Gerlach experiment (see, e.g., [2, 12]), a quantum experiment that generates a truly random sequence.

This random sequence can affect *future* results, so if we know this random sequence, it may help us predict future observations. So, if e can causally influence e', then for some observations x' performed in the small vicinity of e', we have $K(x' \mid r_e) \ll K(x')$.

However, it is clear that this sequence cannot affect the measurement results which are in the *past* (or, more generally, not in the future) of the event e. So, if e cannot causally influence e', then observations x' made in the vicinity of e' are independent on r_e: $K(x' \mid r_e) \approx K(x')$. So, we arrive at the following semi-formal definition:

Definition 3. *For each space-time event e, let r_e denote a random sequence that is generated by an experiment performed in the small vicinity of e. We say that the event e can* causally influence *the event e' if for some observations x' performed in the small vicinity of e', we have*

$$K(x' \mid r_e) \ll K(x').$$

Historical comment. Our definition follows the ideas of casuality as mark transmission [10, 11], with the random sequence as a mark.

Discussion. We have argued that if e *does not* causally influence e', then, no matter what we measure in the vicinity of the event e', we get $K(x' \mid r_e) \approx K(x')$; so, in these cases, the above definition is in accordance with the physical intuition.

On the other hand, if e can causally influence e', this means that we can send a signal from e to e', and as this signal, we send all the bits forming the random sequence r_e. The signal x' received in the vicinity of e' will thus be identical to r_e, so generating x' based on r_e does not require any computations at all: $K(x' \mid r_e) = 0$. Since the sequence $x' = r_e$ is random, we have

$$K(x') \geq \operatorname{len}(x') - C.$$

For a sufficiently long random sequence $r_e = x'$, namely for a sequence for which $\operatorname{len}(x') > 2C$, we have

$$K(x') \geq \operatorname{len}(x') - C > 2C - C = C,$$

so

$$0 = K(x' \,|\, r_e) < K(x') - C$$

and thus,

$$K(x' \,|\, r_e) \ll K(x').$$

So, in these cases, the above definition is also in accordance with the physical intuition.

6.6 Corollary of Our Definition: Space-Time Causality is a Matter of Degree

Our definition of causality uses the notion of randomness: namely, we say that there is a causal re/lation between e and e' if for some random sequence r_e generated in the vicinity of the event e and for measurement results x' produced in the vicinity of e', we have $K(x' \,|\, r_e) < K(x') - C$ for some large integer C.

The larger the integer C, the more confident we are that an event e can causally influence e'. It is therefore reasonable, for each pair of events e and e', to define a *degree of causality c* as the largest integer C for which $K(x' \,|\, r_e) < K(x') - C$. One can check that this largest integer is equal to the difference $c = K(x') - K(x' \,|\, r_e) - 1$. The largest this difference c, the more confident we are that e can influence e'. Thus, this difference can serve as *degree* with which e can influence e'.

In other words, just like randomness turns out to be a matter of degree, causality is also a matter of degree.

Corresponding open problems. It is desirable to explore possible physical meaning of such "degrees of causality": instead of describing the space-time causality, we now have a function $d(e, e')$ that:

- for each pair of events for which e causally precedes e',
- describes to what extent e can influence e'.

Maybe this function $d(e, e')$ is related to relativistic metric – the amount of proper time between e and e'?

Another open problem is related to the fact that he above definition works for localized objects, objects which are located in a small vicinity of one spatial location.

In quantum physics, not all objects are localize in space-time. We can have situations when the states of two spatially separated particles are entangled. It is desirable to extend our definition to such objects as well.

6.7 Conclusions

In this section, we propose a new operationalist definition of causality between space-time events. Namely, to check whether an event e can casually influence an event e', we:

- generate a truly random sequence r_e in the small vicinity of the event e, and
- perform observations in the small vicinity of the event e'.

If some observation results x' (obtained near e') *depend* on the sequence r_e (in the precise sense of dependence described in the section), then we claim that *e can* casually influence e'. On the other hand, if all observation results x' are independent on r_e, then we claim that *e cannot* casually influence e'.

This new definition naturally leads to a conclusion that space-time causality is a matter of degree, a conclusion that is worth physical analysis.

Acknowledgments

References

1. A. Astin, *Are You Smart Enough: How Colleges' Obsession with Smartness Shortchanges Students*, Stylus Publishing, Sterling, Virginia, 2016.
2. R. Feynman, R. Leighton, and M. Sands, *The Feynman Lectures on Physics*, Addison Wesley, Boston, Massachusetts, 2005.
3. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Leaning*, MIT Press, Cambridge, Massachusetts, 2016.
4. V. Kreinovich and A. Ortiz, "Towards a better understanding of space-time causality: Kolmogorov complexity and causality as a matter of degree", *Proceedings of the Joint World Congress of the International Fuzzy Systems Association and Annual Conference of the North American Fuzzy Information Processing Society IFSA/NAFIPS'2013*, Edmonton, Canada, June 24-28, 2013, pp. 1349–1353.
5. M. Li and P. Vitanyi, *An Introduction to Kolmogorov Complexity and Its Applications*, Springer, Berlin, Heidelberg, New York, 2008.
6. J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.
7. C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, W. H. Freeman, New York, 1973.
8. M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, 2000.
9. J. Pearl, *Causality: Models, Reasoning, and Inference*, Cambridge University Press, Cambridge, UK, 2000.

10. S. Psillos, *Causality and Explanation*, McGill-Queen's University Press, Montreal, Kingston, and Ithaca, 2002.
11. W. Salmon, *Causality and Explanation*, Oxford University Press, Oxford, UK, 1997.
12. K. S. Thorne and R. D. Blandford, *Modern Classical Physics: Optics, Fluids, Plasmas, Elasticity, Relativity, and Statistical Physics*, Princeton University Press, Princeton, New Jersey, 2017.
13. L. A. Zadeh, "Fuzzy sets", *Information and Control*, 1965, Vol. 8, pp. 338–353.
14. L. A. Zadeh, "Causality is Undefinable: Toward a Theory of Hierarchical Definability", *Proceedings of the IEEE International Conference on Fuzzy Systems FUZZ-IEEE'2001*, Melbourne, Australia, December 2–5, 2001, pp. 67–68.