2-2004

From Intervals to? Towards a General Description of Validated Uncertainty

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

Gracaliz P. Dimuro

Antonio C. da Rocha Costa

Comments:
UTEP-CS-04-06.
Published as a Technical Report of the Catholic University of Pelotas, Brazil, January 2004.

From Intervals to? Towards a General Description of Validated Uncertainty

Vladik Kreinovich[1], Graçaliz P. Dimuro[2],
and Antônio Carlos da Rocha Costa[2,3]

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

[2]Escola de Informática
Univeridade Católica de Pelotas
96.010-000 Pelotas, RS, Brazil
{rocha,liz}@atlas.ucpel.tche.br

[3]PPGC-UFRGS, Porto Alegre, RS, Brazil

Abstract

In many real-life situations, we are interested in the physical quantities that are difficult or even impossible to measure directly. To estimate the value of such quantity y, we measure the values of auxiliary quantities $x_1, \ldots, x_n$ that are related to y by a known functional relation $y = f(x_1, \ldots, x_n)$, and we then use the results $\widetilde{x}_i$ of measuring x_i to find the desired estimate $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$. Due to measurement errors, the measured values $\widetilde{x}_i$ are slightly different from the actual (unknown) values x_i; as a result, our estimate $\widetilde{y}$ is different from the actual value $y = f(x_1, \ldots, x_n)$ of the desired quantity.

When x_i and y are numbers, then the measurement accuracy can be usually represented in interval terms, and interval computations can be used to estimate the resulting uncertainty in y. In some real-life problems, what we are interested in is more complex than a number. For example, we may be interested in the dependence of the one physical quantity x_1 on another one x_2: we may be interested in how the material strain depends on the applied stress, or in how the temperature depends on a point in 3-D space; in all such cases, what we are interested in is a function. We may be interested in even more complex structures: e.g., in quantum mechanics, measuring instruments are described by operators in a Hilbert space, so if we want to have a precise description of an actual (imperfect) measuring

instrument, what we are interested in is an operator.

For many of such mathematical structures, researchers have developed ways to represent uncertainty, but usually, for each new structure, we have to perform a lot of complex analysis from scratch. It is desirable to come up with a general methodology that would automatically produce a natural description of validated uncertainty for all physically interesting situations (or at least for as many such situations as possible). In this paper, we produce the foundations for such a methodology; it turns out that this problem naturally leads to the technique of *domains* first introduced by D. Scott in the 1970s.

1 Motivation and Outline

The main purpose of this paper is to describe a class of problem for which interval computations are not sufficient, and to design a physically reasonable extension of interval computations that would enable us to solve these problems. To describe these problems, let us first recall why we need interval computations in the first place.

Why computers? We will start this recall by going one step further and recalling why we need *computations* and computers in the first place. In many real-life situations, we are interested in the physical quantities y that are difficult or even impossible to measure directly. For example, there is no easy way to directly measure a distance to a star, the mass of the Earth, or the amount of oil in a given oil field. To estimate the value of such quantity y, we measure the values of auxiliary quantities $x_1, \ldots, x_n$ that are related to y by a known functional relation $y = f(x_1, \ldots, x_n)$. For example, to measure the distance to a star, we can use the parallax method, in which we measure the direction to this star at two different seasons when the Earth is located at the two opposite locations of its solar orbit; to measure the amount of oil, we send ultrasonic signals down the Earth, set sensors at different surface locations, and measure the travel times of these signals from the source to the sensor. The relation f may be an explicit function – as in trigonometric formulas that relate the distance y to star with the angles x_i measured at different seasons; it may be a solution to a system of integro-differential equation as when we reconstruct the geophysical data from the surface measurements, etc. What is necessary is that we have a computer algorithm that, given the values $x_1, \ldots, x_n$, would enable us to reconstruct the desired value y. For convenience, we will denote this algorithm by the same symbol f as the original function.

In the ideal case, once we know the exact values of the quantities x_i, this algorithm will produce the exact value of the desired quantity y. In many real-life situations, this algorithm produces only an approximation to the desired value y.

Once we have this algorithm f, we can then find the estimate for y as follows: if we knew the exact values of x_i, we would be able to apply f to these values and get either the exact desired value y as $y = f(x_1, \ldots, x_n)$ (in the ideal case)

or at least a good approximation to y. In reality, instead of the exact values x_i, we only have the results $\widetilde{x}_1, \ldots, \widetilde{x}_n$ of measuring these auxiliary quantities $x_1, \ldots, x_n$. To compute the desired estimate for y, it is therefore natural to apply the algorithm f to these measurement results. Thus, we arrive at the following estimate for y: $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$.

So, when we cannot measure y directly, we measure it *indirectly*: first, we measure several auxiliary auxiliary quantities $x_1, \ldots, x_n$, and then we apply an algorithm f to the results $\widetilde{x}_i$ of these direct measurements. The algorithm f can be very time-consuming, so time-consuming that we really need a powerful computer to apply it. As a result, such an *indirect measurement*, or *data processing*, is one of the main reasons why we need computers in the first place.

Why interval computations? This explains why we need computations, so why do we need *interval computations*? As we have already mentioned, due to measurement errors, the measured values $\widetilde{x}_i$ are, in general, (slightly) different from the actual (unknown) values x_i: $\widetilde{x}_i \neq x_i$. As a result, even for ideal algorithms, the resulting estimate $\widetilde{y} = f(\widetilde{x}_1, \ldots, \widetilde{x}_n)$ is, in general, different from the actual value $y = f(x_1, \ldots, x_n)$ of the desired quantity. How can we describe this uncertainty?

Before we answer this question, let us mention that there are many different types of uncertainty. In engineering and science, practitioners mainly use *probabilistic* uncertainty, when we describe the probabilities of different values of measurement error $\Delta x_i \stackrel{\text{def}}{=} \widetilde{x}_i - x_i$. In many real-life situations, however, we do not know these probabilities; in many such situations, we only know the upper bounds Δ_i on the absolute value $|\Delta x_i|$ of this measurement error. In such situations, once we have the measurement result $\widetilde{x}_i$, the only information that we have about the actual (unknown) value of the measured quantity x_i is that x_i belongs to the *interval* $\mathbf{x}_i \stackrel{\text{def}}{=} [\widetilde{x}_i - \Delta_i, \widetilde{x}_i + \Delta_i]$. We must therefore transform such interval (validated) uncertainty into the uncertainty of computing y. In other words, given intervals $\mathbf{x}_i$ and an algorithm f, we must find the interval $\mathbf{y}$ of possible values of y, i.e., the interval $\mathbf{y} \stackrel{\text{def}}{=} \{f(x_1, \ldots, x_n) \,|\, x_1 \in \mathbf{x}_1, \ldots, x_n \in \mathbf{x}_n\}$. Techniques for solving this problem are called *interval computations*; see, e.g., [38, 40, 41, 57].

Why go beyond intervals? When we are interested in the value of a number y, then, as we have just mentioned, the measurement accuracy can be usually represented in interval terms, and interval computations can be used to estimate the resulting uncertainty in y. In some real-life problems, however, what we are interested in is more complex than a number.

For example, we may be interested in the dependence of the one physical quantity x_1 on another one x_2: we may be interested in how the material strain depends on the applied stress, or in how the temperature depends on a point in 3-D space; in all such cases, what we are interested in is a function. We may be interested in even more complex structures: e.g., in quantum mechanics, measuring instruments are described by operators in a Hilbert space, so if we want

to have a precise description of an actual (imperfect) measuring instrument, what we are interested in is an operator.

For many of such mathematical structures, researchers have developed ways to represent uncertainty, but usually, for each new structure, we have to perform a lot of complex analysis from scratch. It is desirable to come up with a general methodology that would automatically produce a natural description of validated uncertainty for all physically interesting situations (or at least for as many such situations as possible).

In this paper, we produce foundations for such a methodology. It turns out that this problem naturally leads to a technique of *domains* first introduced by D. Scott in 1970s; see, e.g., [36, 69, 70, 71, 72, 73, 74, 76, 84]. The idea of combining domains with interval uncertainty is not new; see, e.g., [17, 18, 19, 20, 22, 51, 63, 64]. What we will be doing in this paper is explaining why domain techniques are indeed a very natural generalization of interval techniques, and how to make this generalization as general as possible.

Outline. Our main objective is to provide a general technique for describing validated uncertainty. Since we want our description to be as general as possible, we do not want to restrict ourselves to techniques based on real numbers. Instead, we will start with a very general description of a measuring device – so that eventually, from this description, we will be able to construct the real numbers, provide a description of measurements with real numbers, and provide more general descriptions of measured quantities.

As we have already mentioned, a single measuring device provides only an approximate description of a measured quantity. To get a more accurate description of this quantity, we must consider more accurate measuring devices. To get a complete description of a physical quantity, we must therefore consider a (potentially infinite) family of measuring devices that measure this quantity with better and better accuracy. Thus, from the operational viewpoint, a quantity can be associated with the corresponding family of measuring instruments.

Also, since our objective is to describe uncertainty related to data processing (like in interval computations), and data processing means applying an (algorithmically defined) function to the results of direct measurements, we must describe what such functions look like for arbitrary measuring devices.

Data processing means applying a function f, but where does this function f come from? Often, this function has to be experimentally determined, i.e. (to stretch the more traditional meaning of this word), *measured*. A natural question is: how do we describe the procedure ("measuring device") for measuring this function? Once we answer this question, we will be able to describe the (validated) uncertainty associated with measuring a function.

Finally, we will formulate remaining open problems.

2 Measuring Devices: A General Description

2.1 Plan

The main objective of this section is to come up with a general description of validated uncertainty corresponding to a single measuring device. We want to make our description as natural as possible. Because of this, we do not want to simply provide a very general description of a measuring device, with a lot of possible features that may be of use. What we would like to do is to give a step-by-step description of a measuring device, a description in which we provide motivations for each additional feature that we add on each step – motivations that should ideally come from realistic measurement examples.

We start by noticing that every measuring device has only finitely many possible outcomes such as marks on a scale (for an analog device) or digital (usually binary) outputs (for a digital device). Not all the marks may be physically possible: for example, the velocity cannot exceed the speed of light, etc. To distinguish between physically possible and physically impossible marks, we need to have a *theory* that describes both this device and the measured quantity. We will show that, under reasonable assumptions, the existence of such a theory leads to an algorithm for producing a complete list of all physically possible outcomes. This list is the first step in describing a measuring device.

So, we have a set of possible outcomes. On this level, how do we describe uncertainty? Due to measurement uncertainty, when we apply the same measuring device to the same object twice, we may get different measurement results. Some pairs of possible outcomes (x, y) can therefore appear when measuring the same object twice. The information of which pairs can thus occur forms the second step in our description.

In some cases, it is not enough to know which pairs can thus occur; we also need to know which triples, which quadruples, etc. when applying the measuring device three or more times. This additional information constitutes the third step of our description.

Finally, we may need *conditional* statements, i.e., statements of the type "if for some object, we have obtained the values v_1 and v_2, will we be able to get v_3 if we continue measuring the same object?". Such information constitute the fourth and the fifth steps in our description.

We will then show that this is all the validated information that we can extract from a measurement device.

On each step, we will explain how the resulting formalism can be described in terms of the existing mathematical structures.

2.2 First Step: Finite Set of Possible Outcomes

Every measuring device has finitely many possible outcomes. Measuring devices are either analog or digital. For an analog measuring device, the measurement result is a mark on a scale; examples include old-fashioned thermometers, Voltmeters, manometers, scales that measure weight, etc. For

a digital measuring device, the outcome is a sequence of bits (i.e., 0s and 1s); often, for our convenience, this sequence is translated into a decimal number.

For a measuring device with a scale, the measurement result is a mark on the scale, and there are only finitely many marks on the scale.

For a digital measuring device that produces a sequence of bits, there is a bound B on how many bits this device can produce after a single measurement; thus, the overall number of possible measurement results cannot exceed the total number (2^B) of binary sequences of length B.

Not all marks on a scale can be physically possible. Not all the marks (bit sequences) may be physically possible. For example, we may have any velocity marks, but we know, from physics, that the velocity cannot exceed the speed of light. We can have as many temperature marks as we want, but we know, from physics, that the temperature cannot be smaller than the absolute zero (approximately -273°C).

To get a general description of a measuring device, it is desirable to distinguish between physically possible and physically impossible marks.

Comment. As we have mentioned, one of the main objectives of this paper is not simply to provide a new general description of validated uncertainty, but also to provide a motivation for this description. As part of this motivation, we use the common sense words "possible" and "impossible". In this paper, these terms are only used as common sense words: we end up with a formalization of the notions of measuring device, physical quantity, etc., but our analysis does not go as far as producing a formalization of the notion of possibility. In the future, it may be helpful to apply *modal logic* [12, 28, 53, 65] – a part of logic studying terms like "possible" – to also formalize the notions related to possibility.

It is worth mentioning that modal logic have been successfully applied to extend interval approach to uncertainty [4, 8, 30, 31, 44, 46, 80].

Comment: Individual vs. group description of measuring devices. In measurement practice (see, e.g., [62]), there are two ways of describing uncertainty of measuring devices such as sensors.

When sensors are used in critical systems, where the failure can be catastrophic – e.g., in manned space flights – each individual measuring device is calibrated, i.e., we analyze and describe the uncertainty of each individual sensor. For example, when we use three temperature sensors with the same markings produced by the same manufacturer, we calibrate each of these sensors, determining the bias (systematic error) and non-linear distortions characteristic of each individual sensor.

This individual calibration is, however, a very costly procedure. As a result, in most practical applications of measurements, we do not calibrate *individual* sensors; instead, we calibrate the whole collection of mass-produced sensors of

a given time, and use this general description to describe uncertainty of each sensor from this collection.

Our objective is to provide the most general description of validated uncertainty; therefore, we must produce a description that fits both individual and group calibrations. In view of this necessity, in the following text, by a "measuring device", we mean what is calibrated:

- for individual calibrations, we treat each sensor as a measuring device, while
- for group calibrations, by a measuring device we mean the entire collection of mass-produced identically marked sensors.

From this viewpoint, e.g., when we talk about physically possible outcomes, then in the first case, we mean possible outcome of the calibrated sensor, while in the second case, we mean all the outcomes that can appear from using one of the sensors from the sensor collection.

We need a theory. The desired distinction between physically possible and physically impossible marks – i.e., crudely speaking, between physically possible and physically impossible values of the corresponding physical quantity – can only come from physics. In other words, to be able to provide such a distinction, we need a *physical theory*.

We need a theory that also described a measuring device. We need a theory that would describe both the physical quantity (or quantities) measured and/or influenced by the measuring device and the measuring device itself.

Why is it not enough to have a theory that describe the actual physical quantity? Because the measured value depends not only on the quantity but also on a measurement error introduced by the measuring device. For example, if we measure the temperature with a measuring device that has an accuracy 1 degree, then, in principle, it is still physically possible to have the measurement result -274 – when the actual measured temperature is close to the absolute zero value -273; however, the measured value -280 is impossible.

On the other hand, a thermometer with an accuracy ± 10 degrees can produce a reading of -280 – when the actual temperature is close to the absolute zero.

We want a theory that is "full" in some natural sense. Our ultimate objective is to provide solutions to practical problems in which data processing is needed, problems related to engineering and applied science. In engineering and applied science, when we say that there is, say, a theory of flight, this usually means that there is a reasonably full (not partial) description of what can fly and what cannot. There may be some computational difficulties in deciding how well a given complex airplane will fly, difficulties requiring the use of sophisticated supercomputers, but in principle, the problem is solved.

In other words, in engineering and applied science, by a theory, we usually mean a theory that is "full" in some natural sense. This is what we will mean by a theory in this paper.

Comment. It is important to mention that in theoretical physics, the word "theory" also has a different meaning. When we say that, e.g., Einstein produced a theory of gravitation, it does not necessarily mean that his theory is capable of solving all problems related to gravitation. It is known that physical theories often have limitations beyond which they are not applicable; e.g., Einstein's theory of gravitation (General Theory of Relativity) cannot adequately describe the gravitational interaction between quantum objects. In other words, theories in fundamental physics are often *partial*.

There is a reason why in engineering, theories are usually full, while in fundamental science, theories are usually partial: the class of objects that are of interest for an engineering theory is much narrower than the class of objects that is studied by fundamental physics, and it is much easier to come up with a full theory for a smaller class of objects.

A seemingly natural definition of a full theory is not always adequate. At first glance, it may look like a full theory is a one that provides an answer to all relevant questions. This is exactly what is meant by a "decidable theory" in mathematical logic (see, e.g., [5, 23, 49, 66]): a theory in which, for every statement, either this statement or its negation are deducible from the theory.

If we define fullness in this manner, then, for example, if we want to know whether a given mark on a scale of the measuring device is physically possible, then this theory should be able to tell us whether it is physically possible or not.

Unfortunately, this seemingly natural definition does not always capture the physical meaning of the notion of a full theory, the meaning that we are trying to capture in this section. Let us explain the problem with the standard definition on the example of a theory in which there is an upper (or lower) bound on the physical quantity – e.g., the speed of light is the upper bound on velocity, and the absolute zero temperature is the lower bound for temperature.

From the physical viewpoint, when we say that a theory is full, we mean, in particular, that this theory should enable us to compute the exact value of this bound b.

What does it mean to compute b? The bound b is a real number. In the computer, we normally only represent rational numbers, and the actual value b is not necessarily rational. What people usually understand by computing a real number is that for any given accuracy 2^{-k}, we must be able to produce a rational number r_k that is 2^{-k}-close to b; in other words, that we have an algorithm that, given k, produces a rational number r_k for which $|r_k - b| \leq 2^{-k}$. Such algorithms are called *constructive* or *computable* real numbers; see, e.g., [1, 6, 9, 10, 11, 44, 47, 61].

Comment. Since we will be using domains, it is worth mentioning that the notion of a computable real number has been incorporated in domain descriptions; see, e.g., [22].

A full theory should produce an algorithm for computing the bound b, i.e., b should be a computable real number. Similarly, a full theory should compute the exact upper bound Δ on the measurement error of this measuring instrument; hence, Δ should also be a computable real number.

Once we know b and Δ, which marks are physically possible? Without losing generality, let us consider the case of the upper bound. By definition of the measurement error Δx as the difference $\Delta x \stackrel{\text{def}}{=} \widetilde{x} - x$ between the measured and the actual values of the physical quantity, the measured value $\widetilde{x}$ can be represented as the sum of the actual value x and the measurement error Δx. Thus, the largest possible value of $\widetilde{x}$ corresponds to the case when both x and Δ attain their largest possible values. The largest possible value of x is b, and the largest possible value of Δx is Δ. Therefore, the largest possible value of $\widetilde{x}$ is equal to $b + \Delta$.

It is known that the sum of computable real numbers is computable, so $b + \Delta$ is a computable real number. Suppose now that a given mark on scale corresponds to the value v of the measured physical quantity; usually, this value v is a rational number. Then, v physically possible if and only if $v \le b + \Delta$. Thus, to be able to tell whether a given mark on a scale is physically possible or not, we must be able to tell whether a given computable number (in our case, $b + \Delta$) is larger or equal than a given rational number (in our case, v).

Alas, it is known that this problem is, in general, not algorithmically solvable [1, 6, 9, 10, 11, 44, 47, 61]. So, we cannot simply assume that a "full" theory will always tell us which marks are physically possible and which are not. Instead, we must provide a new definition of a full theory and analyze what this definition entails.

Before we do that, let us first describe what we mean by a "theory" in the first place.

What exactly is a theory? In general, in mathematics, to describe a theory means to describe a finite list of its basic statements (called *axioms*) and a finite list of *deduction rules* that enable us to deduce new statements form the ones that we have already proven; see, e.g., [5, 23, 49, 66].

Once we have such a theory, we can algorithmically produce all the statements deducible in this theory: we start with the axioms and apply deduction rules again and again. In this way, we can first produce all the statements that can be derived by a single application of deduction rules, then all the statements that can be produce by two applications of deduction rules, etc. Eventually, each deducible statement will thus be produced.

In theory of computing, a set whose elements can be enumerated by an algorithm is called *recursively enumerable* (r.e., for short). So, when we say that we have a theory, we mean that the set of all statements deduced from this theory is r.e.

R.e. sets S are also called *semi-decidable* (see, e.g., [3, 49]) because the enumerating algorithm provides us with a "semi-decision" procedure for deciding whether a given element x belongs to the set S: We simply generate all the elements of the set S one by one, and compare the given element x with each newly generated element from S. Once we find an element from S that is equal to x, we stop and return the answer "yes". Clearly, if this algorithm returns the answer "yes", it means that $x \in S$; vice versa, if $x \in S$, then this procedure always returns the answer "yes". This procedure is a "semi-decision" procedure because whatever answer it produces is correct, and it is guaranteed to provide a correct answer when the element x actually belongs to S, but it is not required to produce any answer when x is not actually in S.

What kind of statements are we allowing? A physical theory usually describes properties of physically possible states and their transitions: e.g., when we describe the time changes, we must be sure that the energy is preserved, that the moments are preserved, that the overall electric charge is preserved, etc.

If a transition from a given state s into another given state s' is physically *impossible*, we will eventually find it out – by testing all preservation properties.

On the other hand, if the transition is physically *possible* – in the sense that all conserved quantities have the same values in s and s' and therefore, hopefully, there exists an operation that actually transforms s into s' – then we may not be able to deduce this possibility from the theory: there may be infinitely many conserved quantities and during a finite period of time (no matter how long), we can only test finitely many of them, leaving others un-tested.

Similarly, a theory usually describes conditions under which a mathematically defined state is physically possible: the velocity of all the particles cannot exceed the speed of light, the electric charge of each particles must be equal to a whole number of charge quanta, the temperature in each region cannot be smaller than the absolute zero, etc. If a mathematically defined state s is physically *impossible*, we will eventually find it out – by testing all the conditions that a physically possible state must satisfy.

On the other hand, if a mathematically defined state s is physically *possible* – in the sense that all the necessary conditions are satisfied and therefore, hopefully, there exists an operation that actually generates this state s – then we may not be able to deduce this possibility from the theory: there may be infinitely many conditions and and during a finite period of time (no matter how long), we can only test finitely many of them, leaving others un-tested.

In our case, we talk about possible post-measurement states of a measuring device. From this viewpoint, it is reasonable to assume that the set of statements (potentially) covered by a theory includes statements of the type "an outcome v is physically impossible". (It may also happen that a particular theory is able to deduce statements of the type "a mark v is possible" as well, but since we want the most general description, we will not require these statements to be covered by the theory.)

We want the theory to be *correct*; therefore, if the theory implies that an

outcome v is not physically possible, then v should indeed be a physically impossible outcome. Because of this, if a theory implies that an outcomes v is physically impossible, we will call such an outcome *provably impossible*.

As a result, we conclude that *the set of provably impossible outcomes is r.e. (semi-decidable)*.

What exactly is a full theory. Intuitively, a physical theory is full if whatever is not prohibited by this theory actually occurs in real life.

In other words, we require that *if something never happens in real life, the theory must have an explanation for it not happening*, i.e., in precise terms, the theory entails the statement that this particular event never happens.

In other words, if some transformation is never observed, and there is no known law that would prevent this transformation from happening, the physicists usually imply that there must a (yet unknown) additional law that prevent such transformations – i.e., in other words, that the existing theory is not full in the physical sense; see, e.g., [24].

For example, from the viewpoint of the chemical composition, it is perfectly OK for water to dissolve into hydrogen and oxygen: $2H_20 \rightarrow 2H_2 + O_2$; however, if we place water in a bowl, it will never dissolve by itself. A physicist will thus conclude that since this dissolution never happens, there must be an explanation for this – and there actually is such an explanation: the total energy of two water molecules $2H_20$ is smaller than the sum of the energies of two hydrogen molecules $2H_2$ and one oxygen molecule O_2, so this transformation is prohibited by the energy conservation law.

Another example: from the viewpoint of pure particle mechanics, nothing prevents all the air molecules in a room to concentrate in one half of it, leaving vacuum in the other half. However, this never happens. This means that we must have a physical explanation why it never happens, and there is such an explanation – the second law of thermodynamics that prohibits such transformations.

In short, physicists believe that if some event never happens, there must be a theoretical explanation for this event never happening. So, if a theory T provides a full description of a given physical phenomenon, and some event e never happens, this means that the theory T must entail a statement stating that this event e never happens.

How does this apply to measuring devices? Suppose, for example, that we have a measuring device with a scale. Suppose that we use this measuring device again and again and a certain mark never occurs as a result of this measurement. Example: we measure velocity of different particles from the cosmic rays, and the value $v = 400,000$ km/s never occurs. Since this mark never occurs, a full theory T must have an *explanation* for it, i.e., in precise terms, this theory T must imply that this mark never occurs.

The existence of a full theory makes the set of all physically possible outcomes algorithmically listable. Let us show that the existence of a full

theory T enables us to tell, for each outcome v, whether v is physically possible or not.

We want to know whether v will occur in one of the measurements, or whether it never happens.

If the outcome v never happens, then, according to our "definition" of a full theory, T must imply that v never happens. So, if we start with the axioms of the theory T and generate all its conclusions one by one (as we described earlier), we will eventually come up with the statement that v never happens.

If the outcome v will occur, then we can apply the measuring device to all possible objects, and eventually, we will get v. How can we apply the measuring device to all possible objects? We can describe each possible object of measurement by a text in a natural language, a text instructing us how we can find (or generate) this object. We can enumerate all such texts, and apply all such instructions. To be more precise, we can apply the instructions from the first such text for 1 hour, then the instructions from the first two texts for 2 hours, etc. Eventually, we will thus implement each possible object.

So, if we apply the measuring instrument to different objects generated in this way, we eventually get v. (If, by a measuring device, we mean a population of sensors, then we must apply all possible sensors from this population to every object.)

Let us now show how we can check whether v happens or not. We have two algorithmic processes:

- the process that generates all possible statements proven in a theory T, and
- the process that generates all possible objects of measurement.

To check whether an outcome v is possible or not, we simultaneously launch both algorithmic processes and wait until v appears in one of them:

- If an outcome v never occurs (i.e., is physically impossible), then the statement that v never occurs will appear in the first process;
- If an outcome v will occur (i.e., is physically possible), then v will appear as a result of the second process.

Thus, eventually, we will have v as a result of one of these processes – and depending on which one, we will be able to tell whether this outcome v is physically possible or not.

Comment. In mathematical terms, the existence of the first algorithmic process means that the set of all physically *impossible* outcomes is r.e. (semi-decidable). The existence of the second algorithmic process means that the set of all physically *possible* outcomes is also r.e. (semi-decidable). It is well known (see, e.g., [49]) that if a set is r.e. and its complement is r.e., then this set is algorithmically decidable, i.e., there exists an algorithm that checks, for each outcome v, whether this outcome is physically possible or not. This is exactly what we have just shown.

Conclusion: algorithmically listable set of physically possible outcomes. We have shown that, under reasonable assumptions, the existence of a full theory leads to an algorithm for producing a complete list of all physically possible outcomes. This list is the first step in describing a measuring device.

Comment. Since we can always algorithmically check whether each outcome is possible or not, we can always assume that the list of physically possible outcomes is given from the very beginning. Thus, in the following text, we will only consider physically possible outcomes.

Comment. In complex situations related to advanced measurements, we may not have a full theory of the measured quantity. In some such situations, it may still be reasonable to assume that we know the set of all physically possible outcomes, in which case, our description applies. In some other situations, we really do not know which outcomes are physically possible and which are not. How to describe such situations is an important open problem.

Example 1: interval uncertainty. As the first example, we will consider the simplest case of interval uncertainty, when n marks (outcomes) $v_1, \ldots, v_n$ are equally spaced between the smallest value v_1 and the largest value v_n, i.e., $v_2 = v_1 + h$, $v_3 = v_1 + 2h$, $\ldots$, $v_k = v_1 + (k-1) \cdot h$, $\ldots$, $v_n = v_1 + (n-1) \cdot h$, where $h \stackrel{\text{def}}{=} (v_n - v_1)/(n-1)$.

We will also assume that the measurement accuracy Δ is the same throughout the entire scale. For intermediate marks v_k, $1 \le k \le n$, this means that when the outcome is v_k, the actual value of the measured quantity belongs to the interval $S_k \stackrel{\text{def}}{=} [v_k - \Delta, v_k + \Delta] = [v_1 + (k-1) \cdot h - \Delta, v_1 + (k-1) \cdot h + \Delta]$. The lowest mark v_1 means that the actual value is either within Δ from v_1 or smaller than v_1 ("below the scale"), i.e., that the actual value belongs to the set $S_1 \stackrel{\text{def}}{=} (-\infty, v_1 + \Delta]$. Similarly, the highest mark v_n means that the actual value is either within Δ from v_n or larger than v_n ("above the scale"), i.e., that the actual value of the measured quantity belongs to the set $S_n \stackrel{\text{def}}{=} [v_n - \Delta, +\infty)$.

Example 2: counting. An actual counter always has an upper bound n; for binary counters, this upper bound simply comes from the limit on the number of bits that can be represented in this counter. Thus, possible outcomes of counting are: 0, 1, 2, $\ldots$, n, where each number $k < n$ means that we have exactly k units, while the outcome n means that we have either n or more units.

This limitation is similar to how cartoon (and maybe even historic) prehistoric men counted: one, two, many. Here, "one" means exactly one unit, "two" means exactly two units, and "many" means three or more units.

Comment. There exist several different definitions of measurement. From the more engineering viewpoint, a measurement is sometimes defined in a narrower sense, as a procedure for measuring real-valued quantities. From the viewpoint

of this "narrower" definition, counting is not a measurement. However, in more fundamental measurement theory, measurement is usually defined in a broader sense, as any procedure that takes a real object and returns some information about these objects. From this viewpoint, counting is definitely a particular case of measurement.

Since the objective of this paper is to provide a *general* description of measurement-related validated uncertainty, we will use the more general understanding of the word "measurement", the understanding in which procedures like counting are definitely measurements.

Example 3: "yes"-"no" measurements. In many real-life situations, we are not really interested in the actual value of the measured quantity (quantities); all we want to know is whether certain conditions are satisfied. For example, for a single quantity, we may want to know whether the actual value of this quantity exceeds a certain threshold. Such situations are very common in practice: e.g., many actually used measuring devices have a line (often a red line) such that when a measured value (pressure, temperature, etc.) exceeds this line, it is an emergency situation.

In other measuring situations, a situation becomes an emergency if the actual value is below the threshold: e.g., the air pressure in the line that feeds a scuba diver, blood pressure of a patient in an emergency room, etc.

For such measurements, desired answers are "yes" meaning that the tested conditions are satisfied (e.g., the actual value x exceeds the threshold x_0) and "no" meaning that the tested conditions are not satisfied (e.g., the actual value x is still smaller than or equal to x_0). Because of this, such "measurements" can be called *"yes"-"no" measurements.*

In real life:

- when the actual value x of the measured quantity is much smaller than the threshold x_0, we can definitely conclude that $x \le x_0$;
- when the actual value x of the measured quantity is much larger than the threshold x_0, we can definitely conclude that $x > x_0$;
- when the actual value x of the measured quantity is close the threshold x_0, we may not be able to decide whether $x \le x_0$.

Let us illustrate this on the example of interval uncertainty. If we measure a quantity x with accuracy Δ and get the value $\widetilde{x}$ as the result, we can conclude that the actual (unknown) value of x belongs to the interval $[\widetilde{x} - \Delta, \widetilde{x} + \Delta]$. In this case:

- If the largest possible value of x is under the threshold x_0, i.e., if $\widetilde{x} + \Delta \le x_0$, we can guarantee that the actual value x does not exceed x_0.
- If the smallest possible value of x is above the threshold x_0, i.e., if $\widetilde{x} - \Delta > x_0$, we can guarantee that the actual value x exceeds x_0.

- In the remaining cases, when x_0 belongs to the interval $[\widetilde{x} - \Delta, \widetilde{x} + \Delta)$ of possible values, we can have both $x \le x_0$ and $x > x_0$, so we cannot decide whether $x \le x_0$.

For "yes"-"no" measurements, we therefore have three possible outcomes: "yes", "no", and "unknown". In the computer, "yes" is usually represented as 1 and "no" as 0. In logic, the value "unknown" is often abbreviated as U. So, the set of all physically possible outcomes is $\{0, 1, U\}$.

Comment. If the measured quantity can take all possible real values, then all three outcomes are physically possible. However, in some physical situations, not all real values are physically possible and thus, not all outcomes of a "yes"-"no" measurement may be physically possible.

For example, if all physically possible values of x are bounded by a bound x_B, and x_0 exceeds this bound, then the outcome 1 is physically impossible. (We may still get U if the actual value is close to x_B.)

This is a clear example of when, because of our knowledge of the underlying physics, the set of outcomes that are physically possible in this particular measuring situation is smaller than the set of all outcomes that are potentially possible.

Another example of such a situation is when we measure a quantized quantity such as electric charge that can only take values q, $2q$, ..., $n \cdot q$, etc., and the measurement error Δ is smaller than a half of this quantum, then, based on the measurement results, we can always decide whether $x \le x_0$. Indeed, in this case, the interval $[\widetilde{x} - \Delta, \widetilde{x} + \Delta]$ of width $2\Delta < q$ can contain only one value of the type $k \cdot q$ and therefore, from the measurement, we can uniquely determine the actual value x of the measured electric charge. In such cases, the outcome U is physically impossible, and the only physically possible outcomes are 0 and 1.

Yet another example of a situation when not all outcomes are physically possible emerges when we consider *repeated* "yes"-"no" measurements.

Example 3a: repeated "yes"-"no" measurements. If, based on a single measurement, we cannot decide whether we exceeded the threshold, it is reasonable to repeat the measurement with a more accurate instrument. We may not want to incorporate this more accurate measurement instrument into our measuring device.

One may ask: if we already have the more accurate instrument available, why not use this more accurate instrument in the first place? Why do we need to use the less accurate instrument at all?

The answer to this question is that the more accurate instrument usually requires more time (and more other resources such as maintenance cost, energy consumption, etc.) to operate. It is therefore reasonable to use this more expensive-to-use instrument only if necessary. This is what, e.g., medical doctors do if they suspect a disease: first, they order a simple test like a simple blood test or a mammogram. If this simple test returns the answer "no", this means

that a patient does not have a suspected disease. If the test returns the answer "yes", e.g., if certain parameters of the blood test are clearly way above the norm, this often means that the disease is confirmed. In some cases, however, based on the results of the simple test, we cannot make a definite decision: all we know that there is something suspicious, so the doctors order a more sophisticated test.

If a more sophisticated (and more expensive-to-use) test still does not provide us with a definite answer, we may order an even more accurate (and even more expensive-to-use) test, etc.

What are the possible outcomes of such repeated measurements? In the first measurement, we get the values "yes", "no", and "unknown". To distinguish between the results of the first (less accurate) and the second (more accurate) measuring instruments, we will denote the outcomes of the first measuring instrument by 0_1, 1_1, and U_1; the outcomes of the second measuring instrument will be denoted by 0_2, 1_2, and U_2.

If the result of the first measurement is 0_1 or 1_1, we stop; otherwise, we perform the second measurement whose result is also either "yes", "no", or "unknown. The outcomes of the second measuring instrument will be denoted by 0_2, 1_2, and U_2. As a result, we get the following outcomes: 0_1, 1_1, $U_1 0_2$, $U_1 1_2$, and $U_1 U_2$.

Comment. In principle, to get more accurate results, we could always apply both measuring instruments. In this case, we would have the following 5 possible outcomes: $0_1 0_2$, $1_1 1_2$, $U_1 0_2$, $U_1 1_2$, and $U_1 U_2$.

If we did not know that the two measurements measure exactly the same thing, then we could also have the outcomes $0_1 1_2$ and $1_1 0_2$.

If we did not know that the second measuring instrument is more accurate than the first one, then we could also have outcomes $0_1 U_2$ and $1_1 U_2$.

Our knowledge about the measured quantity and the measuring device enables us to see that out of $9 = 3^2$ mathematically possible pairs of the values 0, 1, U, only 5 are physically possible outcomes.

Example 4: a combination of several independent measuring instruments. In the previous example, we considered repeated measurements of the same quantity. This is an example of *dependent* measurements, when after knowing the result of the first measurement, we can restrict the possible results of the second measurement.

In many real-life situations, a measuring device consists of several *independent* measuring instruments that measure independent quantities. For example, a meteorological station measure temperature, pressure, and wind velocity; these values are independent of each other.

Suppose that a measuring device consists of m measuring instruments, and we know that these measuring instruments are *independent* in the sense that once we know the results of some of the measurement, it does not change which outcomes are possible for other measurements. Let X_1 be the set of all physically

possible outcomes of the first measurement, let X_2 be the set of all physically possible outcomes of the second measurement, ..., and let let X_m be the set of all physically possible outcomes of the m-th measurement. An outcome of this measuring device is a sequence of outcomes of these m individual measurement: $x = (x_1, \ldots, x_m)$ where $x_1 \in X_1$, ..., $x_m \in X_m$. Since we assume that the measurements are independent, all such tuples are possible, hence the set of all possible outcomes of a measuring device is the set of all such tuples, i.e., the Cartesian product $X_1 \times \ldots \times X_m$.

2.3 Second Step: Pairs of Compatible Outcomes

How do we describe uncertainty: main idea. So, we have a set of possible outcomes. On this level, how do we describe uncertainty? Due to measurement uncertainty, when we apply the same measuring device to the same object twice, we may get different measurement results.

This is a known fact for anyone who has ever measured anything with a real measuring device, be it current or voltage or temperature.

For example, suppose that we measure temperature by using a measuring device with an accuracy of ± 2 degrees. If the actual temperature is, say, 36.6, then, due to measurement errors, the measurement result can take any value from $36.6 - 2 = 34.6$ to $36.6 + 2 = 38.6$. If this measuring instrument has marks corresponding to $0, 1, 2, \ldots, 100$ degrees, then for the actual temperature of 36.6, the only possible outcomes are the ones that lie within the interval $[34.6, 38.6]$, i.e., the marks 35, 36, 37, and 38.

The measurement error usually has a random component. As a result, when we repeatedly measure the temperature of the same object (with the actual temperature of 36.6), we get, in general, different measurement results in different measurements: we may get 35 in the first measurement, 37 in the second one, etc.

Let us show how this uncertainty-related possibility can be described in the most general case.

Comment on quantum measurements. Before we proceed with the description of uncertainty via repeated measurements, it is important to mention that the above text is only applicable to non-quantum measurements.

In terms of measurements, the main difference between non-quantum and quantum physics is as follows: Of course, every measurement involves interaction with a measured object and thus, affects this object. For example, one way to measure the temperature of a hot object is to attach a thermometer to this object; since the original temperature of the thermometer was smaller than the temperature that we measure, we thus slightly cool down the object.

In non-quantum physics, we can make the effect of the measuring device on an object as small as possible, so when we want to measure the same quantity again, the actual value will not change much. This is usually done by making the measuring devices much smaller than the objects.

Quantum physics describes very small objects, so in quantum physics, we usually measure properties of very small objects by using much larger measuring instruments. As a result, every quantum measurement drastically changes the properties of a measured object – to the extent that often, the measured object is completely destroyed in the process of measurement. Moreover, there is a theoretical limit on how much we can restrict the effect of measurement on an object – the limit known as Heisenberg's Uncertainty Principle.

For quantum objects, we thus often cannot literally apply the same measuring device twice to the same object – and even if we can, we will get a completely different result. This is not so bad because many microscopic objects measured in quantum measurements have a short lifetime anyway. When we measure properties of quantum objects, we usually do not mean measuring properties of a perishable short-living individual object; rather, when we say that we are interested in the mass of a Hydrogen atom or in a magnetic moment of an electron, we mean that we have a generator that generates multiple copies of the same (or almost the same) object. In such situations, instead of considering repeated measurements of the same object, we should therefore talk about repeated measurements of different objects generated by the same object generator.

In the following text, for simplicity, we will mainly talk about repeated measurement of the same object, but what we will mean is that either we have multiple measurements of the same object, or we have multiple measurement performed on different copies of the object – copies produced by the same object generator.

Some pairs of outcomes are compatible (close), some are not. We have already mentioned that due to measurement uncertainty, multiple measurement of the same object can lead to different outcomes.

For example, if we measure temperature with the accuracy of 2 degrees, and possible outcomes are 0, 1, 2, etc., then the outcome 0 can happen if the actual temperature is in the interval $[-2, 2]$; the outcome 1 can happen if the actual temperature is in the interval $[-1, 3]$, etc.

When the outcomes are very different, e.g., 0 and 5, the corresponding intervals $[-2, 2]$ and $[3, 7]$ do not intersect – which means that there is no actual value for which repeated measurements would produce 0 and 5.

When the outcomes are close, e.g., 1 and 3, then the corresponding intervals $[-1, 3]$ and $[1, 5]$ have a non-empty intersection – which means that there are actual values for which repeated measurements can produce 1 and 3 – namely, all the values from this intersection.

It is therefore natural to say that the physically possible outcomes x and y are *compatible*, or *close* if for some object, multiple measurements can lead to x and y. We will denote compatible outcomes by $\sim$. For example, for the above thermometer, $1 \sim 3$ but $0 \not\sim 5$.

We can also say that in this case, the outcomes $x \sim y$ are *indistinguishable* by the given measuring device: if we got x as an outcome for one object, and y

as an outcome for another object, then we cannot guarantee that these objects are different: they may be identical.

The existence of a full theory makes the set of all compatible pairs of outcomes algorithmically listable. How do we tell which pairs of outcomes are compatible and which are not? As in the first step, the only way to distinguish between compatible and non-compatible pairs is to have a physical theory that describes both the measured quantity and the measuring device.

Similarly to the first step, we can argue that this theory should be able to make predictions about incompatibility of different pairs. We can also argue that this theory should be full in the sense that if some pair (x, y) never occurs in a repeated measurement, then the theory should have an explanation for it never occurring, i.e., the theory should entail that this pair never occurs. If it does occur, then we will be able to see it in one of the actual measurements.

Similarly to the first step, we can therefore conclude that the set of all compatible pairs is decidable (algorithmically listable).

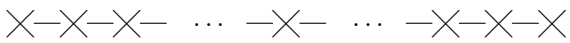
Conclusion: algorithmically listable set of compatible pairs of outcomes. We have shown that, under reasonable assumptions, the existence of a full theory leads to an algorithm for producing a complete list of all compatible pairs of outcomes. Thus, at the current (second) step of describing a measuring device, we can describe it as a pair $\langle X, \sim\rangle$, where X is a finite set (of all physically possible outcomes), and $\sim$ is a symmetric and reflexive relation on X describing which outcomes are compatible.

Description in terms of existing mathematical structures. From the mathematical viewpoint, a pair $\langle X, \sim\rangle$ is nothing else but a finite *graph*.

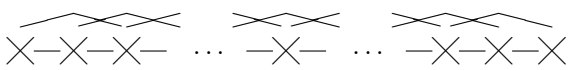
In general, a graph in which vertices correspond to sets and two vertices are connected by an edge if and only if the corresponding sets intersect is called an *intersection graph* [35, 52]. For the case when sets and intervals, the intersection graph is called an *interval graph*; see, e.g., [27].

Comment. Since we will be dealing with domains, it is worth mentioning that in domain theory, graphs – also called *webs* – form the basis of *binary Girard domains*; see, e.g., [32, 33, 34, 78]. From this viewpoint, on our current (second) step we describe measuring devices by using binary Girard domains. Such a domain description has been actually used in our previous papers to describe interval uncertainty; see, e.g., [19, 20].

Example 1: interval uncertainty. For the simplest interval uncertainty, as we have mentioned, there exists a constant C such that v_i and v_j are compatible if and only if $|i - j| \le C$. Let us show how the corresponding graphs look like. For $C = 1$, each point is only connected to its immediate neighbors, so the graph takes the following form:



For $C = 2$, each element is compatible not only with its immediate neighbors, but also with immediate neighbors of immediate neighbors. This results in the following graph:



Example 2: counting. Let us recall that for counting with an upper bound n, we have $n+1$ different outcomes $0, 1, \ldots, n$, such that each outcome $k \le n-1$ means that we have exactly k objects, and the outcome n means that we have n or more objects.

Based on this meaning, we can easily conclude that the outcome is uniquely determined by the actual number of objects. Thus, no matter how many objects we actually have, repeated measurement will always produce exactly the same outcome. Conclusion: no two different outcomes are compatible, i.e., we have a *completely disconnected* graph:

$$0 \quad 1 \qquad \cdots \qquad k \qquad \cdots \qquad (n-1) \quad n$$

Comment. It is worth mentioning that in quantum physics, the situation with counting is completely different: we may have a physical state that is a superposition (in some precise sense) of states with, say, 2 and 3 particles. Thus, when we measure the number of particles in such a state, we may get 2 in the first measurement and then 3 in the second measurement. In other words, the graph corresponding to quantum counting is different – no longer completely disconnected.

Example 3: "yes"-"no" measurements. In the case of "yes"-"no" measurements, we have three possible outcomes: 0, 1, and U ("unknown").

- If the actual value is 0, we can get 0 and U;
- if the actual value is 1, we can get 1 and U.

Thus, 0 and U are compatible; 1 and U are compatible; but 0 and 1 are, of course, incompatible. So, we end up with the following graph:

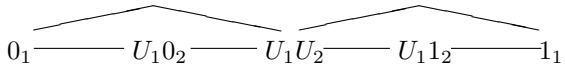
$$0 \text{———————} U \text{———————} 1$$

Example 3a: repeated "yes"-"no" measurements. In the case of repeated "yes"-"no" measurements, we have five possible outcomes: 0_1, 1_1, $U_1 0_2$, $U_1 1_2$, and $U_1 U_2$.

- If the actual value is 0, we can get 0_1, $U_1 0_2$, and $U_1 U_2$;

- if the actual value is 1, we can get 1_1, $U_1 1_2$, and $U_1 U_2$.

Thus, 0_1, $U_1 0_2$, and $U_1 U_2$ are compatible with each other; also 1_1, $U_1 1_2$, and $U_1 U_2$ are compatible; but values corresponding to 0 and 1 are, of course, incompatible. So, we end up with the following graph:



Example 4: a combination of several independent measuring instruments. When a measuring device consists of m independent measuring measurements, then the outcome x of the device is a tuple $(x_1, \ldots, x_m)$ of outcomes corresponding to individual measuring instruments. Since the measurements are independent, the two tuples $x = (x_1, \ldots, x_m)$ and $y = (y_1, \ldots, y_m)$ are compatible if for each k, the corresponding measuring results are compatible. Thus, if we have m graphs $\langle X_i, \sim_i \rangle$ corresponding to m different measuring instruments, then the measuring device can be described by a graph $\langle X, \sim \rangle$, where $X = X_1 \times \ldots \times X_m$ and

$$x = (x_1, \ldots, x_m) \sim y = (y_1, \ldots, y_m) \leftrightarrow x_1 \sim_1 y_1 \,\&\, \ldots \,\&\, x_m \sim_m y_m.$$

Thus defined graph is called a *Cartesian product* of graphs $\langle X_1, \sim_1 \rangle$, $\ldots$, $\langle X_m, \sim_m \rangle$.

Computational complexity of the graph representation of a measuring device: general case. Our ultimate objective is to provide a description of a measuring device that would be helpful in indirect measurements and data processing, i.e., that would be helpful in computing.

From this viewpoint, when selecting a description of a measuring device, we must keep in mind that the intent of this description is that it be used in future computations. In particular, we must make sure that the description of a measuring device does not take up too many bits of computer memory – because when processing this description, we must access each bit at least once, and if there are many of them, the processing will take long.

In our case, a measuring device is represented by a graph. Let us recall how many bits we need to store a generic graph. Let us denote the total number of vertices of a graph by n. Vertices of the graph can be denoted by $v_1, \ldots, v_n$. To describe a graph, we must know, for each i from 1 to n and for each j from 1 to n (that is different from i), whether the vertices v_i and v_j are compatible. For each pairs (i, j), we need 1 bit to store the corresponding information. There are $n(n-1)/2$ pairs (i, j) with $i \neq j$. Therefore, we need $n(n-1)/2 = O(n^2)$ bits to store the information about a general graph.

Is it feasible? For a real-life measuring device, the accuracy may be 0.1% of the overall scale, which means that there are about $n \approx 10^3$ possible outcomes. If we want to describe the list of all possible outcomes, then we need n bits. If we also want to describe a graph structure, then we need $n^2 \approx 10^6$ bits. This amount of computer memory is feasible.

The above amount of computer memory is feasible but it is, for large n, much larger than the number of bits that are need to simply store all possible outcomes. Can we do better?

Computational complexity of the graph representation of a measuring device: case of the simplest interval uncertainty. To see whether we can do better, let us start with the above simplest case of interval uncertainty. In this case, we have an equally spaced scale with outcomes v_1, $v_2 = v_1 + h$, $v_3 = x_1 + 2h$, $\ldots$, $v_k = v_1 + (k-1) \cdot h$, $\ldots$, $v_n = v_1 + (n-1) \cdot h$, and the measurement accuracy Δ is the same throughout the entire scale. For $k = 2, \ldots, n-1$, if the outcome is v_k, this means that the actual value of the measured quantity belongs to the interval $S_k = [v_k - \Delta, v_k + \Delta] = [v_1 + (k-1) \cdot h - \Delta, v_1 + (k_1) \cdot h + \Delta]$. The outcome v_1 means that $x \in S_1 = (-\infty, v_1 + \Delta]$, and the outcome v_n means that $x \in S_n = [v_n - \Delta, +\infty)$.

In this case, two outcomes v_i and v_j are compatible if the corresponding intervals intersect. One can easily see that this intersection is non-empty if $v_i - \Delta \le v_j + \Delta$ and $v_j - \Delta \le v_i + \Delta$, i.e., substituting the above expressions for $v_i = v_1 + (i-1) \cdot h$ and $v_j = x_1 + (j-1) \cdot h$ into these inequalities, that $|i - j| \cdot h \le 2\Delta$, i.e., $|i - j| \le 2 \cdot \Delta / h$.

In other words, *if the outcomes v_i and v_j are compatible, then the corresponding indices i and j should be C-close to each other* (where $C \stackrel{\text{def}}{=} 2\Delta/h$).

In this case, once we fix h and Δ (and thus, C), we do not need $O(n^2)$ bits to describe which outcome is compatible to which: it is sufficient, for each i from 1 to n, to describe which outcomes within a C-neighborhood of i are compatible with i. Since C is fixed, we need a constant number of bits to store this information for each i. Thus, the overall amount of computer memory that is needed to store this information about all the inputs is $O(n)$.

Comment. In the simplest interval case, there are n outcomes corresponding to 1, 2, $\ldots$, n, and the two outcomes are compatible if and only if $|i - j| \le C$. In this case, we can decrease the required amount of computer memory even further, because all we need to store are the integers n and $C \le n$. How much space do we need to store these two numbers? By using b bits, we can store all the integers from $0 \ldots 0$ (b bits) $= 0$ to $1 \ldots 1$ (b bits) $= 2^b - 1$; so, to store a number n, we need b bits where $2^{b-1} - 1 < n \le 2^b - 1$, i.e., $b \approx \log_2(n)$ bits. Storing an integer $C \le n$ requires even fewer bits, so overall, we need $O(\log_2(n))$ bits.

However, in the general interval case, as we will see, we will need all n bits.

Computational complexity of the graph representation of a measuring device: general case of interval uncertainty. The above example was the simplest case of interval uncertainty. In general, the outcomes may not be equally spaced: there may be more marks per unit in some part of the scale, and fewer in other parts of the scale. Also, the measurement accuracy Δ_i may

be different for different parts of the measurement scale. Let us show that the above conclusion holds for the general case of interval uncertainty as well.

Let $\Delta = \max \Delta_i$ denote the largest radius of the interval, and let $h = \min(v_{i+1} - v_i)$ denote the smallest possible distance between the sequential outcomes. For every $i < j$, we have $v_j - v_i = (v_j - v_{j-1}) + (v_{j-1} - v_{j-2}) + \ldots + (v_{i+1} - v_i)$. Each of $j - i$ terms in this sum cannot be smaller than h, hence $v_j - v_i \le (j - i) \cdot h$. In general, $|v_i - v_j| \ge |i - j| \cdot h$.

If the two outcomes v_i and v_j are compatible, then the corresponding intervals intersect, hence $v_i - \Delta_i \le v_j + \Delta_j$ and $v_j - \Delta_j \le v_i + \Delta_i$. From these two inequalities, we can conclude that $v_i - v_j \le \Delta_i + \Delta_j$ and $v_j - v_i \le \Delta_i + \Delta_j$, hence $|v_i - v_j| \le \Delta_i + \Delta_j$. We know that Δ is the largest of the values Δ_i, so $\Delta_i + \Delta_j \le 2\Delta$. On the other hand, we have proven that $|v_i - v_j| \ge |i - j| \cdot h$. Thus, we can conclude that $|i - j| \cdot h \le 2\Delta$, i.e., that $|i - j| \le C \stackrel{\text{def}}{=} 2\Delta/h$.

So, once we fix h and Δ (and thus, C), *we can use only* $O(n)$ *bits* to store the information about which physically outcomes are compatible with which.

Computational complexity of the graph representation of a measuring device: lower bound for the case of the general interval uncertainty. Let us show that in the general interval case – in contrast to to the simplest interval case – we do, in general, need at least $c \cdot n$ bits (actually, at least $n/2$ bits) to store the graph information.

To prove this, we will first show that there are at least $2^{n/2}$ different graphs, at least as many as there are bit sequences of length $n/2$. Namely, to each such bit sequence $b_1 \ldots b_{n/2}$, we put into correspondence the following case of interval uncertainty: $v_i = i$, $\Delta_1 = \Delta_3 = \Delta_5 = \ldots = 1$, and $\Delta_{2i} = 1 + b_i$. Here, $h = 1$, $\Delta = \max \Delta_i = 2$, hence $C = 2\Delta/h = 4$.

In this case, the interval corresponding to an odd outcome v_{2i+1} is $S_{2i+1} = [2i, 2i + 2]$, and the interval corresponding to an even outcome v_{2i} is $S_{2i} = [2i - 1 - b_i, 2i + 1 + b_i]$. Let us show that, based on the intersection graph, we can uniquely determine the original sequence b_i. Indeed, the intervals corresponding to $2i$ and $2i + 3$ (or $2i - 3$) intersect if and only if $2i + 1 + b_i \ge 2i + 2$, i.e., if and only if $b_i \ge 1$. Since b_i can take only values 0 and 1, the outcomes $2i$ and $2i + 3$ are compatible if and only if $b_i = 1$. Thus, based on the intersection graph, we can uniquely reconstruct the binary sequence b_i. So, $2^{n/2}$ different binary sequences lead to different intersection graphs.

Hence, there are at least $2^{n/2}$ different intersection graphs. If we use b bits to store the information about a graph, then we can have at most 2^b different representations. Thus, to be able to distinguish between $2^{n/2}$ different graphs, we need at least $n/2$ bits. The statement is proven.

Computational complexity of the graph representation of a measuring device: case of multi-D uncertainty. In the case of multi-D uncertainty, each outcome v_i can be represented as a point in the corresponding m-dimensional space R^m, and the related uncertainty can be described by a set

$S_i \subseteq R^m$ containing v_i. Thus, when the measurement device produces an outcome v_i, the only information that we have about the actual (unknown) value $x \in R^m$ of the measured m-dimensional quantity is that $x \in S_i$.

In the case of interval uncertainty, the set S_i is an m-dimensional box $[\underline{x}_1, \overline{x}_1] \times \ldots \times [\underline{x}_n, \overline{x}_n]$. In many practical situations, the uncertainty is best described by a ball or, more generally, an ellipsoid, etc.

Let us show that in all such cases, we do not need to use all $O(n^2)$ bits to store the information. For each set S_i, we can define its *radius* Δ_i as the largest possible distance $d(v_i, s)$ between v_i and points $s \in S_i$: $\Delta_i = \sup\{d(v_i, s) \,|\, s \in S_i\}$. Let Δ denote the largest of these radii: $\Delta = \max \Delta_i$.

How close are the points v_i? For each point v_i and for each distance d_0, we can denote the *point density* $\rho(v_i, d_0)$ as the ratio $N(v_i, d_0)/V(d_0)$, where $N(v_i, d_0)$ is the overall number of points $v_j \neq v_i$ in the d_0-vicinity $B_{d_0}(v_i) = \{x \,|\, d(x, v_i) \leq d_0\}$ of the point v_i, and $V(d_0)$ is the volume of this vicinity. For each point v_i, we can then define the largest point density $\rho(v_i)$ as $\max \rho(v_i, d_0)$, where the maximum is taken over all positive real numbers d_0. One can easily see that the maximum is attained when d_0 is equal to the distance between v_i and one of the points $v_j \neq v_i$ – because when we increase d_0 without hitting one of these distance, the value $N(v_i, d_0)$ remains the same but the volume increases. Thus, we can alternatively define $\rho(v_i)$ as $\max_{j \neq i} \rho(v_i, d(v_i, v_j))$.

We can then define the *configuration density* ρ as the largest of the n values $\rho(v_i)$.

Let us fix the configuration radius Δ and the configuration density ρ. If the sets S_i and S_j corresponding to two different outcomes v_i and v_j intersect, this means that there exists a point $s \in S_i \cap S_j$. For this point, $d(s, v_i) \leq \Delta_i$ and $d(s, v_j) \leq \Delta_j$, hence $d(v_i, v_j) \leq d(v_i, s) + d(s, v_j) \leq \Delta_i + \Delta_j$. Since Δ is the largest of the values Δ_i we can conclude that $d(v_i, v_j) \leq 2\Delta$.

Thus, if the two outcomes v_i and v_j are compatible, we conclude that the corresponding points in an m-D space are 2Δ-close. We know, from the definition of the configuration density, that in the 2Δ-vicinity of the point v_i, there are $\leq C \stackrel{\text{def}}{=} \rho \cdot V(2\Delta)$ different points $v_j \neq v_i$. Thus, each outcome v_i can have no more than C compatible outcomes. When Δ and ρ are fixed, C is a constant.

Let us show that because of this property, we only need $O(n \cdot \log(n))$ bits to store the graph information. Indeed, it is sufficient to store, for each outcome v_i, the list of all its neighbors (= compatible outcomes). Each neighbor is described by a number from 1 to n. We already know that to store such a number, we need $log_2(n)$ bits; thus, to store th information about all C neighbors of a given outcome, we need $C \cdot \log(n)$ bits. To store the information about the entire graph, we need to store this information for each of n outcomes v_i. Thus, overall, *we need* $n \cdot (C \cdot \log(n)) = O(n \cdot \log(n))$ *bits.*

Computational complexity of the graph representation of a measuring device: general case of localized uncertainty. Let us now consider the general case of localized uncertainty.

Let M be the set of all possible values of a measured quantity. For 1-

D measurements, $M = R$; for multi-D measurements, $M = R^m$; in general, as we have mentioned, M can be the set of all the functions o the set of all operators from quantum mechanics. Each outcome v_i can be represented as a point in the set M. The related uncertainty can be described by a set $S_i \subseteq M$ containing v_i. Thus, when the measurement device produces an outcome v_i, the only information that we have about the actual (unknown) value $x \in M$ of the measured quantity is that $x \in S_i$.

Intuitively, "localized" means that there is a natural definition of closeness (i.e., a metric) on the set M, so all the points from the uncertainty set S_i are reasonably close, and in each vicinity of each point v_i, there are not too many other points $v_j \neq v_i$.

Similarly to the multi-D case, for each set S_i, we can define its *radius* Δ_i as the largest possible distance $d(v_i, s)$ between v_i and points $s \in S_i$: $\Delta_i = \sup\{d(v_i, s) \,|\, s \in S_i\}$. Let Δ denote the largest of these radii: $\Delta = \max \Delta_i$.

Lt us also assume that there is a computable function $N(d_0)$ such that for each point v_i and for each distance d_0, there are no more than $N(d_0)$ points $v_j \neq v_i$ in the d_0-vicinity of v_i. (Based on the values v_i, we can determine $N(d_0)$ as $\max_i N_i(d_0)$, where $N_i(d_0)$ is the total number of points $v_j \neq v_i$ in the d_0-vicinity of the point v_i.)

Let us fix the configuration radius Δ and the "density function" $N(d_0)$. If the sets S_i and S_j corresponding to two different outcomes v_i and v_j intersect, then, as in the multi-D case, we can conclude that $d(v_i, v_j) \leq 2\Delta$. Thus, if the two outcomes v_i and v_j are compatible, the corresponding points in M are 2Δ-close. We know, from the definition of the density function $N(d_0)$, that in the 2Δ-vicinity of the point v_i, there are $\leq C \stackrel{\text{def}}{=} N(2\Delta)$ different points $v_j \neq v_i$. Thus, each outcome v_i can have no more than C compatible outcomes. When Δ and $N(d_0)$ are fixed, C is a constant.

Thus, as in the multi-D case, *we only need $O(n \cdot \log(n))$ bits to store the corresponding graph information.*

2.4 Third Step: Subsets of Compatible Outcomes

From pairs to subsets. Instead of repeating a measurement twice, we can repeat the same measurement three and more times. As a result, we may get three or more different outcomes for the same object. To describe this situation, we can use the same word "compatible", and say that the outcomes $v_1, \ldots, v_m$ are *compatible* if for some object, as a result of repeated measurement, we get these values v_i (in whatever order). So, to get a better description of a measuring instrument, we must not only describe which pairs are compatible, we must also describe what triples, what quadruples, etc., are compatible.

Is information about compatible pairs sufficient? If we only know which pairs are compatible, can we then decide which subsets are compatible?

First, we should mention that once we know all compatible pairs, some subsets can be easily dismissed as not compatible. If a set of outcomes $v_1, \ldots, v_m$

is compatible, i.e., if there is an object for which repeated measurements lead to $v_1, \ldots, v_m$, then, of course, each pair (v_i, v_j) is also compatible. So, if two elements from a set are not compatible with each other, this means that the set itself is not compatible either.

A natural question is whether the opposite is also true: if we have a finite set $\{v_1, \ldots, v_m\}$ in which every two outcomes are compatible, if this set compatible?

Information about compatible pairs is sufficient for intervals. Let us show that for interval uncertainty, the answer to the above question is "yes". Indeed, in interval uncertainty, an outcome $\widetilde{x} = v_i$ means that the actual value of the measured quantity belongs to an interval $S_i = [\underline{x}_i, \overline{x}_i]$.

For example, if all the measurements are equally accurate, then $\underline{x}_i = v_i - \Delta$ and $\overline{x}_i = v_i + \Delta$. In these terms, two outcomes v_i and v_j are compatible if the corresponding two intervals have a non-empty intersection $S_i \cap S_j \neq \emptyset$, and a set $\{v_1, \ldots, v_m\}$ of outcomes is compatible if $S_1 \cap \ldots \cap S_m \neq \emptyset$.

What happens if every two intervals intersect? Two intervals $[\underline{x}_i, \overline{x}_i]$ and $[\underline{x}_j, \overline{x}_j]$ have a non-empty intersection if and only if $\underline{x}_i \leq \overline{x}_j$ and $\underline{x}_j \leq \overline{x}_i$. Thus, if every two intervals intersect, this means that for every i and j, we have $\underline{x}_i \leq \overline{x}_j$ – i.e., every lower endpoint is not larger than every upper endpoint. We will now show that the intersection of all m intervals is non-empty by showing that this intersection contains the largest of the lower endpoints $x \stackrel{\text{def}}{=} \max \underline{x}_i$. In other words, we will show that this value x belongs to each of m intervals $[\underline{x}_j, \overline{x}_j]$, i.e., that $\underline{x}_j \leq x \leq \overline{x}_j$. Indeed:

- since x is the largest of the lower endpoints, it is larger than or equal to each of them, in particular, $\underline{x}_j \leq x$.
- On the other hand, since x is the largest of m real numbers, it is equal to one of them: $x = \underline{x}_k$ for some k. We know that $\underline{x}_k \leq \overline{x}_j$ so $x \leq \overline{x}_j$.

The statement is proven.

Information about compatible pairs is not sufficient in the general case. We will now show that for a general non-interval uncertainty, the information about which pairs are compatible and which are not is not sufficient to determine which sets are compatible.

When the measured quantity is 1-D, i.e., its possible values are real numbers, then intervals describe the measurement uncertainty. In many real-life situations, we measure a multi-D quantity $x = (x_1, \ldots, x_m)$, i.e., the quantity that consists of several real-valued components $x_1, \ldots, x_m$. For example, in meteorology, we may want to measure different coordinates of the wind vector. Intervals can describe the uncertainty resulting from measuring each individual component. If the components are physically independent, then the resulting multi-D uncertainty of measuring the entire m-dimensional quantity x is described by a multi-D box.

In real life, different components x_i are often highly dependent; as a result, after each measurement, the set of possible values $(x_1, \ldots, x_m)$ of the measured

m-dimensional quantity is different from a box. Often, ellipsoids provide a good descriptions for this uncertainty.

Ellipsoids naturally come from probabilistic uncertainty: if the measurement error is described by a multi-D Gaussian distribution, then the confidence set – described as the set of all the values for which the probability density exceeds a certain threshold – is an ellipsoid. Ellipsoids have also been successfully used to describe validated uncertainty; see, e.g., [7, 13, 14, 25, 26, 29, 67, 68, 75, 79]. Not only ellipsoids work well; it has been experimentally shown that in many practical situations, ellipsoids work better than other families of sets [13, 14]. Moreover, it was theoretically proven that under certain reasonable conditions, ellipsoids indeed form the best family [26, 50].

For ellipsoids, it is easy to find an example when every two ellipsoids have a non-empty intersection, but the three ellipsoids have no common point; see Fig. 1.

Figure 1: Case when each pair of outcomes is compatible but the three outcomes are not compatible

Thus, if these ellipsoids correspond to outcomes of a measuring device, we get an example where every two outcomes are compatible, but the entire set of outcomes is not compatible. So, in general, in addition to describing which *pairs* of outcomes are compatible, we must also describe which *sets* of outcomes are compatible.

The existence of a full theory makes the family of all compatible sets of outcomes algorithmically listable. How do we tell which sets of outcomes are compatible and which are not? Similarly to the pairwise comparison, the only way to distinguish between compatible and non-compatible sets is to have a physical theory that describes both the measured quantity and the measuring device.

Again, we can argue that this theory should be able to make predictions about incompatibility of different sets. We can also argue that this theory should be full in the sense that if some set S never occurs in a repeated measurement, then the theory should have an explanation for it never occurring, i.e., the theory should entail that this set never occurs. If it does occur, then we will be able to see it in one of the actual measurements.

Similarly to the first step, we can therefore conclude that the set of all compatible sets is decidable (algorithmically listable).

Conclusion: algorithmically listable family of compatible sets of outcomes. We have shown that, under reasonable assumptions, the existence of a full theory leads to an algorithm for producing a complete list of all compatible sets of outcomes. Thus, at the current (third) step of describing a measuring device, we can describe it as a pair $\langle X, \mathcal{S}\rangle$, where X is a finite set (of all physically possible outcomes), and $\mathcal{S} \subseteq 2^X$ is a subset of 2^X with the following two properties:

- *subset property*: if $S \in \mathcal{S}$ and $S' \subseteq S$, then $S' \in \mathcal{S}$;
- *singleton property*: for every $x \in X$, then $\{x\} \in \mathcal{S}$.

Such a description was first proposed and developed in [17, 18].

Comment. The subset condition comes from the above observation that a subset of a compatible set is always compatible. The singleton condition simply means that every outcome from the set X is physically possible – i.e., that X is the set of all physically possible outcomes.

Description in terms of existing mathematical structures: simplicial complexes. In algebraic topology, the above structure is known as a simplicial complex; see, e.g., [48]. Simplicial complexes are a technique for describing polytope-type topological spaces. A simplicial complex is made of *simplexes*.

The basic 1-dimensional simplex is a straight line segment v_1v_2 connecting two points v_1 and v_2. This segment has two vertices: x_1 and x_2.

A 2-dimensional simplex is a triangle $v_1v_2v_3$ – a convex hull of a three-point set $\{v_1, v_2, v_3\}$. The triangle has three vertices: v_1, v_2, and v_3. The triangle $v_1v_2v_3$ has three faces – 1-dimensional simplexes v_1v_2, v_2v_3, and v_1v_3.

A 3-dimensional simplex $v_1v_2v_3v_4$ is a tetrahedron – a convex hull of the four-point set $\{v_1, v_2, v_3, v_4\}$. A tetrahedron has four vertices: v_1, v_2, v_3, and v_4. A tetrahedron has 4 faces $v_1v_2v_3$, $v_1v_2v_4$, $v_1v_3v_4$, and $v_2v_3v_4$; each of these faces is a triangle – i.e., a 2-dimensional simplex, etc.

To this, we can add a 0-dimensional simplex x, which is simply a point (and its own vertex).

A simplicial complex is a finite collection of simplexes; these simplexes may share faces with each other. This complex represent a *union* of the corresponding sets, i.e., the set of all the points that belong to at least one of the simplexes from the complex. From the set-theoretic viewpoint, a face is a subset of a simplex. So, if a simplex S belongs to a given complex C, and F is its face, then it make sense to conclude that F also belong to the same complex – because all the points from F automatically belong to S and thus, belong to the union representing the complex C.

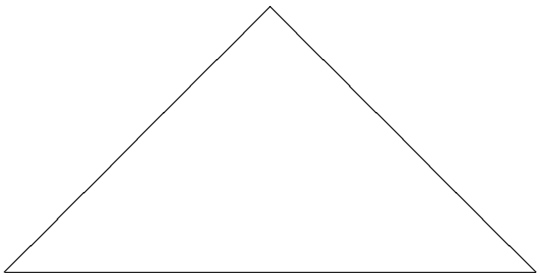
In view of this observation, a simplicial complex is usually defined as a finite collection of simplexes such that if a simplex belongs to this collection, then all its faces also belong to it.

As we have mentioned, each simplex can be uniquely described by its vertices. Thus, if we denote the set of all the vertices of all the simplexes by X, we can alternatively define a simplicial complex as a collection of subsets of X.

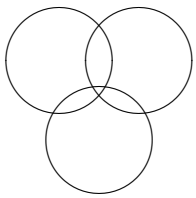
How will the face property be represented in these terms? A face of a simplex defined by a set of k vertices can be represented by a subset of $k-1$ elements. If we require that all the faces are included in C, then we conclude that faces of faces, faces of faces of faces, etc., are also included. In set terms, it means that with every simplex S (= subset of X), every subset of S also belongs to the complex. Thus, a simplicial complex is a collection of sets with a subset property – exactly what we need for our description of a measuring device.

Resulting geometric representation of a measuring device. The similarity between our description of a measuring device and the notion of a simplicial complex enables us to provide a geometric representation of a measuring device.

For example, suppose that we consider two different measuring devices. In both devices, there are 3 physically possible outcomes v_1, v_2, and v_3, and in both devices, every pair of outcomes is compatible. So, from the graph viewpoint, when we only represent pairwise compatibility of pairs of physically possible outcomes, both devices are represented by the same graph:

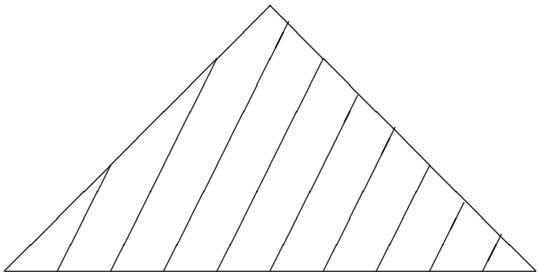


The difference between the two measuring devices – that we want to describe geometrically – is that in one of them (as in Fig. 1), the three values v_1, v_2, and v_3 are incompatible, while in the second device, the three values v_i are compatible:



From the geometric viewpoint, the first measuring device only has three 1-D

simplexes and no 2-D ones, so it is represented by the above hollow triangle. The second measuring device also has a 2-D simplex $\{v_1, v_2, v_3\}$, so it is represented as a *filled* triangle:



Towards description in terms of existing mathematical structures: domains. Another natural description of the structure $\langle X, \mathcal{S}\rangle$ is a description in terms of *domains* [36, 69, 70, 71, 72, 73, 74, 76, 84].

The notion of a domain was originally invented by D. Scott to describe the different amounts of information that we may have about a given object or class of objects. The main elements of a domain are these different pieces of information, and the main relation is the relation "x contains all the information contained in y" (and maybe some more information), the relation that is usually denoted as $x \sqsupseteq y$.

In this case, we also say that in comparison with a more complete piece of information x, y provides less information, i.e., y *approximates* x; this relation is denoted by $y \sqsubseteq x$.

By the very meaning of these notions, the approximation relation is reflexive, transitive, and antisymmetric, i.e., it is a partial order relation.

The approximation relation can be naturally reformulated in terms of a subset relation: indeed, to each piece of information x, we can put into correspondence the set S_x of all possible situations that are consistent with this information. In these terms, if x contains all the information contained in y, this means that every situation consistent with x is also consistent with y, i.e., S_x is a subset of S_y. In other words, $x \sqsupseteq y$ means that $S_x \subseteq S_y$.

If we acquire two pieces of information x (corresponding to the set S_x) and y (corresponding to the set S_y), then what is the resulting set of possible situations? If we know both x and y, this means that we know that the situation must satisfy all the conditions imposed by the knowledge x, and it must also satisfy all the conditions imposed by the knowledge y. The set of all situations that satisfy conditions imposed by x is S_x; similarly, the set of all conditions imposed by y is S_y. Thus, we conclude that the actual situation must belong to the set $S_x \cap S_y$.

Comment. It should be mentioned that the pieces of information x and y may be inconsistent, e.g., x may mean that the actual value is positive, while y

means that the actual value is negative. In this case, when no situation can simultaneously satisfy x and y, the intersection $S_x \cap S_y$ is an empty set.

So, the family of sets S_x corresponding to different pieces of information x must be closed under intersection. In terms of the partial order $\sqsupseteq$ ($=\subseteq$), intersection can be defined as the least upper bound; this least upper bound is usually denoted by $x \sqcup y$.

Comment. In domain theory, there are additional considerations related to the *infinite* families of pieces of information. At any given state of the knowledge, however, we only deal with *finitely* many possible pieces of information – for one reason that the overall number of symbols is limited, and there are only finitely many words expressed by that many symbols.

If we restrict ourselves to *finite* families of pieces of information, then the formal definition of a domain can be presented as follows (see, e.g., [73]): a *closure system* is a non-empty family of sets closed under intersection, and a *domain* is a partially ordered set isomorphic to a closure system. This is exactly what we have described above.

Comment. The above definition of a domain in terms of closure systems (to be more precise, a more complex definition covering infinite families as well) was used by D. Scott in his earlier papers. From the viewpoint of understanding the basic motivations this is a very good and clear definition – and this is why we used it here.

However, from the viewpoint of mathematical analysis of domains, this is a very indirect and inconvenient definition: a domain is a partially ordered set, in our case, *finite* partially ordered set, but it is defined in terms of a representation as families of (possibly *infinite*) sets. From this viewpoint, it is desirable to use equivalent definitions that define a domain directly in terms of the properties of the partial order. Such equivalent definitions do exist, and they are the ones that are used as definitions in most books on domains and their applications.

Comment. It is also worth mentioning that usually, the case of inconsistent information (corresponding to the combination of two inconsistent pieces of information) is omitted from the description of a domain, so we will, in most cases, omit it too.

On the other hand, the situation when we have not performed any measurement yet – and thus, we have no information about the measured quantity – is usually included in the description of the domain. This complete absence of information is the smallest amount of information that we can possible have, so it is usually denoted by a "bottom" symbol $\bot$.

How to reformulate the above description of a measuring device in terms of domains? Let us consider a measuring device $\langle X, \mathcal{S}\rangle$ with the set X of physically possible outcomes. Domain theory describes pieces of information,

so, in order to describe the measuring device in terms of domain theory, we must find out what information about the object do we get after we use this device to measure an object.

After a single measurement, we get a single measurement result $x \in X$. So, after this measurement, the only information that we gain about the actual value of the measured quantity is that this value is consistent with x. It is natural to use this same outcome x to represent these pieces of information.

After repeated measurements, we can get several different outcomes. For example, if in two different measurements, we get, e.g., x and x', this means that the actual value of the measured quantity is consistent both with x and with x'. For every set $S \subseteq X$ of outcomes, it is also natural to use this same set as a notation for this piece of information.

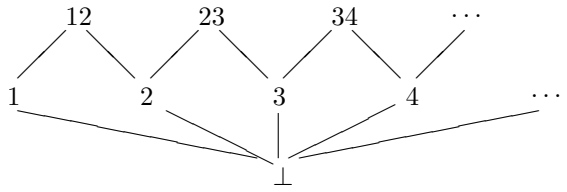
When we have both outcomes x and x', we thus clearly have more information about the measured object than if we had only one of them: $xx' \sqsupseteq x$ and $xx' \sqsupseteq x'$. From the set-theoretic viewpoint, the set of the values that are compatible with both outcomes x and x' is an intersection of the sets corresponding to x and x'; in domain terms, it is the least upper bound: $xx' = x \sqcup x'$.

In these terms, the structure $\langle X, \mathcal{S}\rangle$ is naturally described as a domain.

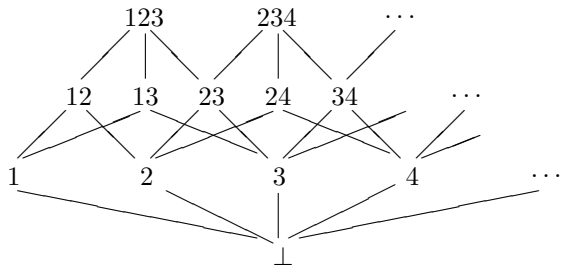
Comment. In domain theory, structures of the type $\langle X, \mathcal{S}\rangle$, where $\mathcal{S} \subseteq 2^X$ satisfies the subset property and the singleton property, form the basis of *Girard domains*; see, e.g., [32, 33, 34, 78]. From this viewpoint, on our current (third) step we describe measuring devices by using Girard domains. This observation was first explicitly mentioned in [17, 18].

Example 1: interval uncertainty. For the simplest interval uncertainty, as we have mentioned, there exists a constant C such that v_i and v_j are compatible if and only if $|i - j| \le C$.

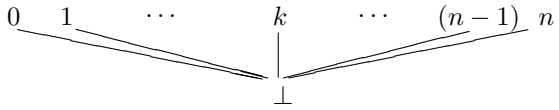
For $C = 1$, each point is only connected to its immediate neighbors, and no three outcomes are compatible. So, we get the following domain:



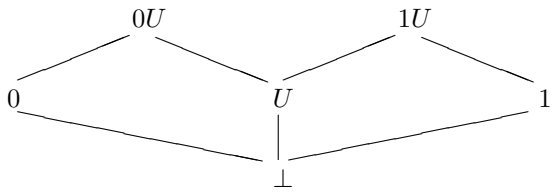
For $C = 2$, each element is compatible not only with its immediate neighbors, but also with immediate neighbors of immediate neighbors. Here, every triple $(i, i + 1, i + 2)$ is compatible; no quadruples are compatible. This results in the following domain:



Example 2: counting. For counting, we have $n+1$ possible outcomes 0, 1, ..., n; every two of them are incompatible. Thus, the corresponding domain has the following form:



Example 3: "yes"-"no" measurements. For "yes"-"no" measurements, the outcomes 0 and U are compatible, and the values 1 and U are compatible, but 0 and 1 are not compatible.Thus, we arrive at the following domain:



Comment. It should be mentioned that the value "unknown" (U) is *different* from the case of complete absence of information $\bot$. Indeed, $\bot$ means that we have not performed any measurement, so we have no information yet.

The value U means that we have performed a measurement, but this measurement did not enable us to decide whether, e.g., the desired quantity x exceeds the given threshold x_0. The fact that we did not gain any information on whether $x > x_0$ does not necessarily mean, however, that we gained no information at all. For example, if our measuring process consists of measuring x with accuracy Δ and comparing the result $\widetilde{x}$ with x_0, then, as we mentioned earlier, the fact that after the measurement, we still cannot decide whether $x > x_0$ means that $\widetilde{x} - \Delta \le x_0 < \widetilde{x} + \Delta$, i.e., that $x_0 - \Delta < \widetilde{x} \le x_0 + \Delta$.

Since we know that the actual value x of the measured quantity is Δ-close to $\widetilde{x}$, i.e., $|x - \widetilde{x}| \le \Delta$, we can thus conclude that $x \le \widetilde{x} + \Delta \le (x_0 + \Delta) + \Delta = x_0 + 2\Delta$ and similarly, that $x \ge \widetilde{x} - \Delta > (x_0 - \Delta) - \Delta = x_0 - 2\Delta$, i.e., that $x_0 - 2\Delta < x \le x_0 + 2\Delta$.

In this example, although we did not decide whether the answer to the desired question is "yes" or "no", we did gain some information about the actual value x that we did not know before – and thus, in this case, U has more information than $\bot$: $U \sqsupseteq \bot$ and $U \neq \bot$.

Example 4: a combination of several independent measuring instruments. When a measuring device consists of m independent measuring instruments, then the outcome x of the device is a tuple $(x_1, \ldots, x_m)$ of outcomes corresponding to individual measuring instruments. Since the measurements are independent, a set of tuples $S \subseteq X_1 \times \ldots \times X_m$ is compatible if for each k, the corresponding measuring results are compatible.

How can we describe this in formal terms? For each k and for each tuple $x = (x_1, \ldots, x_k, \ldots, x_m)$, we can define *$k$-th projection* $\pi_k(x)$ as of the k-th component of the tuple x. For every set S of tuples, we can define $\pi_k(S)$ as the set of k-th projections of all the tuples from S: $\pi_k(S) \stackrel{\rm def}{=} \{\pi_k(x) \,|\, x \in S\}$.

In these terms, if we have m simplicial complexes $\langle X_i, \mathcal{S}_i\rangle$ corresponding to m different measuring instruments, then the measuring device can be described by a simplicial complex $\langle X, \mathcal{S}\rangle$, where $X = X_1 \times \ldots \times X_m$ and

$$S \in \mathcal{S} \leftrightarrow \pi_1(S) \in \mathcal{S}_1 \,\&\, \ldots \,\&\, \pi_m(S) \in \mathcal{S}_m.$$

Thus defined simplicial complex will be called a *Cartesian product* of the simplicial complexes $\langle X_1, \mathcal{S}_1\rangle$, $\ldots$, $\langle X_m, \mathcal{S}_m\rangle$.

Computational complexity of the simplicial complex representation of a measuring device: a general case. How many bits do we need to store the new information about the measuring device? Let n denote the number of physically possible outcomes. A natural way to describe a domain structure $\mathcal{S}$ is to describe, for each of 2^n subsets $S \subseteq X$, whether S belongs to $\mathcal{S}$.

For each of 2^n sets $S \subseteq X$, we need one bit to store this information, so overall, it is sufficient to use 2^n bits.

This representation of a domain is excessive because, due to the subset property, once we know that $S \in \mathcal{S}$, we already know that every subset $S' \subseteq S$ also belongs to $\mathcal{S}$. As a result of this excess, it is possible to use fewer bits to store the domain information. It turns out, however, that we cannot decrease the required storage space too much: we will show that to describe a general domain, we cannot use fewer than $\approx 2^n$ bits.

To prove it, we will show that we can have at least $\approx 2^{2^n}$ different possible domain structures $\mathcal{S}$ on a set X of n elements. Let $\mathcal{H}$ denote the class of all subsets of X that have exactly $\lfloor n/2 \rfloor$ elements; such sets will be called *half-size*. For each subset $H \subseteq \mathcal{H}$, we can determine $S(H)$ as the family of all sets $S \subseteq X$ that are contained in one of the half-size sets $h \in H$. This family clearly has both subset and singleton properties.

Let us show that different sets H lead to different domains $S(H)$. Specifically, we will show that once we know $S(H)$, we can determine H uniquely – as the set of all half-size sets from $S(H)$. Indeed, clearly, every half-size set from

H also belongs to $S(H)$. Vice versa, suppose that a half-size set h belongs to $S(H)$. By definition, it means that h is a subset of a half-size set $h' \in H$. Since a proper subset always has fewer elements that the original set, a half-size set h cannot be a proper subset of another half-size set h'; thus, we can only have $h = h'$. Since $h' \in H$, we conclude that $h \in H$.

So, we can construct at least as many different domains as there are subsets in the set of all half-sets. In other words, if we denote, by A, the total amount of all half-size subsets of X, then we we can construct at least 2^A different domains.

The number A of different subsets of half-size is equal to $\binom{n}{\lfloor n/2 \rfloor}$. The number of combinations $\binom{a}{b}$ is known to be equal to $\dfrac{a!}{b! \cdot (a-b)!}$, and it is known that asymptotically, $n! \approx (n/e)^n$. Therefore,

$$A = \binom{n}{\lfloor n/2 \rfloor} \approx \frac{n!}{(n/2)!^2} \approx \frac{(n/e)^n}{\left((n/2e)^{n/2}\right)^2} = \frac{(n/e)^n}{(n/2e)^n} = 2^n.$$

So, we can have at least $2^A \approx 2^{2^n}$ different domain structures.

We have already mentioned, in analyzing computational complexity of the graph representation, that if we use b bits to store the information about a structure, then we can have at most 2^b different structures. Thus, to be able to distinguish between 2^{2^n} different domain structures, we need at least 2^n bits.

So, *in general we need 2^n bits to store the information about a domain structure.*

This amount is no longer feasible: as we have mentioned, for a real-life measuring device, we an have about $n \approx 10^3$ possible outcomes. If we want to describe the list of all possible outcomes, then we need n bits. If we also want to describe a generic Girard domain structure, then we need $2^n \approx 2^{10^3}$ bits – this amount is larger than the overall number of particles in the Universe.

It is therefore extremely important to find cases when we can use fewer bits. It turns out that interval and, in general, localized uncertainty provides exactly such cases.

Computational complexity of the simplicial complex representation of a measuring device: case of interval uncertainty. For interval uncertainty, as we have mentioned, every outcome v is only compatible with neighbors that are C-close to v. Thus, every compatible set containing i is a subset of the set $\{i-C, i-C+1, \ldots, i-1, i, i+1, \ldots, i+C\}$ consisting of $2C+1$ elements. For each i from 1 to n, there are finitely many ($\le 2^{2C+1}$) such subsets, so in the interval case, we need $O(n)$ bits to store the entire Girard domain information.

Computational complexity of the simplicial complex representation of a measuring device: case of multi-D uncertainty. In most applications, multi-D uncertainty is described by *bounded convex* sets $S_i \subseteq R^m$. Bounded convex sets satisfy the known *Helly theorem* (see, e.g., [15, 21, 35, 52, 54, 55]):

if every $m+1$ subsets of a finite family of sets have a common point, then the entire family has a non-empty intersection.

In this case, we do not need to list all the intersecting (compatible) subsets $S \subseteq \mathcal{S}$: only subsets that have no more than $m+1$ outcomes. There are no more than $\begin{pmatrix} n \\ m+1 \end{pmatrix} = O(n^{m+1})$ such subsets, so we need at most $O(n^{m+1})$ bits to store the information on which of these subsets belong to $\mathcal{S}$ and which do not.

Comment. In the particular case $m = 1$, convex sets $S \subseteq R^1$ are intervals, so Helly theorem becomes the above-mentioned property of intervals and pair-wise intersections. It is worth mentioning that *boxes* $[\underline{x}_1, \overline{x}_1] \times \ldots \times [\underline{x}_m, \overline{x}_m]$ in the multi-D space also have this *2-Helly property* that if every two boxes from a finite family intersect, then the entire family has a non-zero intersection. Moreover, as shown in [54, 55], the family of all the boxes is, in some reasonable sense, the only family of convex sets that has this 2-Helly property.

Computational complexity of the simplicial complex representation of a measuring device: general case of localized uncertainty. In the case localized uncertainty, each outcome has no more than C neighbors. To store the information about these neighbors, as we have mentioned in the graph case, we need $O(n \cdot \log(n))$ bits.

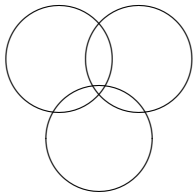
For each outcome v, every compatible set containing this outcome is a subset of the set of v's neighbors. For each of n outcomes v, there are finitely many ($\leq 2^C$) such subsets, so in the localized case, we need $O(n)$ bits to store the subset information.

Overall, we need $O(n \cdot \log(n)) + O(n) = O(n \cdot \log(n))$ bits – the same amount as in the graph case.

2.5 Fourth Step: Conditional Statements about Possible Outcomes

Subsets of compatible outcomes do not always give a complete description of a measuring device. In the previous subsection, we have shown that subsets of compatible outcomes often provide more information about a measuring device than pairs of compatible outcomes – and thus provide a *more* complete description of a measuring device. However, as we will see, this description is still not always fully complete.

To show this incompleteness, let us consider two different measuring devices each of which has three possible outcomes v_1, v_2, and v_3. In the first device, we have interval uncertainty: v_1 corresponds to the interval $S_1 = [1-2, 1+2] = [-1, 3]$, v_2 corresponds to the interval $S_2 = [2-2, 2+2] = [0, 4]$, and v_3 corresponds to the interval $S_3 = [3-2, 3+2] = [1, 5]$. In the second device, the uncertainty corresponding to each outcome is described by a circle:



In both devices, the three sets corresponding to the outcomes v_i have a common point, hence all three outcomes are compatible. So, from the viewpoint of subsets of compatible outcomes, the two measuring devices are described by exactly the same structure $\langle X, \mathcal{S}\rangle$, where $X = \{v_1, v_2, v_3\}$ and

$$\mathcal{S} = \{\{v_1\}, \{v_2\}, \{v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_1, v_2, v_3\}\}.$$

However, from the physical viewpoint, there is a difference between these devices, a difference that is not captured by the above common description.

Namely, for the second (circle-related) measuring device, there are values that belong to both v_1 and v_3 but do not belong to v_2. This means that it is possible to have an object for which v_1 and v_3 are possible outcomes, but for which we will never get v_2 as a possible outcome.

In contrast, for the first (interval-related) measuring device, the intersection of v_1 and v_3 is contained in v_2. Thus, when for some object, repeated measurement produced v_1 and v_3, we are sure that further repeated measurement will eventually produce the outcome v_2 – if not for this particular device, but at least for one of the devices characterized by the same uncertainty.

What we do we need to add to the subsets description to capture the missing information about a measuring device? We have shown that there is a difference between the measuring devices that is not captured by a subset description. How can we describe such a difference in general terms?

In the first (interval-related) measuring device, for every object, if v_1 and v_3 are possible outcomes, then v_2 is also a possible outcome. In the second measuring device, there exist objects for which v_1 and v_3 are possible outcomes but v_2 will never occur.

The subset information formalizes statements describing which outcomes are possible for an object, statement of the type

"For some object, v_1 and v_3 are possible outcomes."

To capture the above difference, we must also formalize *conditional* statements of this type, i.e., statements of the type

"If for some object, v_1 and v_3 are possible outcomes, then for this same object, v_2 is also a possible outcome."

Comment. Up to this subsection, we have been describing formalisms that have already been proposed and used to describe uncertainty of measuring devices. Many of our motivations described in the previous subsections are new, many results are new, but the main definition has already been described in the previous papers. From this point on, we start describing a *new* approach to the description of measuring devices. This approach is not mathematically new – it uses existing mathematical formalisms, what is new is their application to measuring devices.

The existence of a full theory makes the set of all true conditional statements algorithmically listable: an argument. How do we tell which conditional statements are true for a given measuring device and which ar not?

As before, the only way to distinguish between conditional statements that are true for this device and conditional statements that are not true is to have a physical theory that describes both the measured quantity and the measuring device.

Once we have such a theory, how can we tell, e.g., whether after measuring v_1 and v_3, we will always be able to get v_2?

We will show that, as before, the existence of a full theory makes the set of all true conditional statements algorithmically listable. The argument will be similar to the previous arguments, but slightly more complicated.

First, we remark that for every object for which we have observed v_1 and v_3, there are two options: either v_2 never occurs, or v_2 occurs at some future moment of time – whether by measuring with this particular measuring instrument or by measuring with some other measuring instrument that was made by using the same specifications and the same technology (and thus, belongs to the same collection comprising this measuring device).

If v_2 never occurs, then, as before, it is reasonable to expect that a full theory T provide an explanation for it never occurring; in other words, the statement that v_2 never happens must be entailed from T.

If v_2 actually occurs, we will eventually observe it by applying this (or similar) measuring instruments to this object.

Thus, as before, we can simultaneously launch two algorithmic processes:

- the process of deriving all possible logical consequences of the theory T; and
- the process of applying this (and similar) measuring devices to the same object.

Eventually, we will get v_2 in one of these two processes:

- if v_2 never occurs, then the first process will produce a statement that it never occurs (a statement derivable from the theory T);
- if v_2 occurs at some moment of time, the second process will produce v_2 as a measurement result.

Thus, similarly to the previous cases, for each object for which v_1 and v_3 both occur, we will know whether v_2 will occur or not.

In other words, for each object o for which we have observed v_1 and v_3, we can algorithmically decide whether v_2 never happens, i.e., we can algorithmically decide whether the statement $P(o) \stackrel{\text{def}}{=} \forall t\,(M(o,t) \neq v_2)$ is true, where $M(o,t)$ means the result of measuring the object o at time t.

In general, for some objects o for which v_1 and v_3 have been observed, this property $P(o)$ holds, for some other objects o, this property $P(o)$ does not hold.

If the property $P(o)$ never holds, then it is reasonable to expect that a full theory T should provide us with an explanation for this not occurring, i.e., the statement that $P(o)$ never holds should follow from the theory T.

On the other hand, if the property $P(o)$ holds for some object o, then, by testing all possible objects, we will eventually find the object o for which this property is true.

Thus, we can simultaneously launch two new algorithmic processes:

- the process of deriving all possible logical consequences of the theory T; and
- the process of testing the property $P(o)$ for all possible objects o.

Eventually, we will get $P(o)$ in one of these two processes:

- if $P(o)$ never holds, then the first process will produce a statement that it never holds (a statement derivable from the theory T);
- if $P(o)$ holds for some object o, the second process will produce this object.

Thus, we will know whether there exists an object o for which $P(o)$ holds.

According to the definition of the property $P(o)$, if this property holds for some object o, then the original conditional statement is not valid; if this property does not hold for any object o, then the implication expressed by our conditional statement is valid.

So, we can algorithmically decide whether a given conditional statement holds or not.

Family of conditional statements: natural properties. We have shown that, under reasonable assumptions, the existence of a full theory leads to an algorithm for producing not only a complete list $\mathcal{S}$ of all compatible sets of outcomes, but also a complete list of all true conditional statements of the type

"If for some object, all outcomes from a set $S \subseteq X$ have been observed, then an outcome $v \notin S$ will also be eventually observed for this object."

Each of the statements and the entire family of such statements must satisfy the following natural properties.

First, an individual conditional statement makes sense only if $S \cup \{v\} \in \mathcal{S}$ – otherwise, by definition of $\mathcal{S}$, once S has been observed, we cannot observe v

for the same object. Once the family $\mathcal{S}$ is fixed, we can say that a pair $\langle S, v\rangle$, where $S \in \mathcal{S}$, is a *possible condition* if $v \notin S$ and $S \cup \{v\} \in \mathcal{S}$.

Second, if S implies v, and we have observed $S' \supseteq S$, this means that we have, in particular, observed S, so S' also implies v.

Third,

- if S implies v, meaning that whenever all outcomes from S are observed for some object, then v will be eventually observed, and
- $S \cup \{v\}$ implies v', meaning that once we observed S and v, we will eventually observe v',

then we can conclude that once we observe S, we will eventually observe v', i.e., that S implies v'.

Thus, we arrive at the following conclusion.

Conclusion: algorithmically listable family of conditional statements. We conclude that at the current (fourth) step of describing a measuring device, we can describe it as a triple $\langle X, \mathcal{S}, \mathcal{C}\rangle$, where:

- X is a finite set (of all physically possible outcomes),
- $\mathcal{S} \subseteq 2^X$ is a subset of 2^X that satisfies the *subset property* (if $S \in \mathcal{S}$ and $S' \subseteq S$, then $S' \in \mathcal{S}$) and the *singleton property* ($\mathcal{S}$ includes all singletons), and
- $\mathcal{C}$ is a family of pairs $\langle S, v\rangle$, where $S \subseteq X$ and $v \in X$, that satisfies the following three properties:
  - Every pair from $\mathcal{C}$ is a *possible condition*, i.e., $v \notin S$ and $S \cup \{v\} \in \mathcal{S}$.
  - *superset property*: If $\langle S, v\rangle \in \mathcal{C}$, $S' \supseteq S$ and $\langle S', v\rangle$ is a possible condition, then $\langle S', v\rangle \in \mathcal{C}$; and
  - *transitivity*: if $\langle S, v\rangle \in \mathcal{C}$ and $\langle S \cup \{v\}, v'\rangle \in \mathcal{C}$, then $\langle S, v'\rangle \in \mathcal{C}$.

Description in terms of existing mathematical structures: deduction relation. From the mathematical viewpoint, the set $\mathcal{C}$ of pairs $\langle S, v\rangle$ has properties similar to the properties of the *deduction relation* $\vdash$, i.e., the relation $S \vdash v$ meaning that the statement v follows from the set of statements included in the set S; see, e.g., [5, 23, 49, 66].

Indeed, if S implies v, then S' also implies v – hence we have a superset property. If $S \vdash v$ and $S, v \vdash v'$, then $S \vdash v'$ – hence we have transitivity.

Comment. There are two main differences between our description and the traditional deduction relation:

- First, we only consider situations in which the set S of statements is consistent, while in mathematical logic, the deduction relation also covers the case when S is inconsistent. For inconsistent sets S, in classical logic, $S \vdash v$ for all possible statements v, so we can easily extend our relation to inconsistent sets S as well.
- Second, in classical logic, we also consider the cases when $v \in S$: in this case, of course, $S \vdash v$. These cases are also easy to add to our relation.

In summary, if $\vdash$ is our relation, then we can define the classical deduction relation $\vdash'$ as follows:

$$S \vdash' v \leftrightarrow (S \vdash v) \vee v \in S \vee (S \text{ is inconsistent}).$$

Vice versa, if we have a classical logical deduction relation $\vdash'$, we can define the notion "S is consistent" as $S \not\vdash' False$ and then define the consistency-bounded relation $\vdash$ as follows:

$$S \vdash v \leftrightarrow (S \vdash' v) \,\&\, v \notin S \,\&\, (S \text{ is consistent}).$$

From the purely mathematical viewpoint, as we can see, our modification of the traditional definition of the deduction relation is equivalent to the traditional definition of deduction relation. The only reason why we modified the traditional definition is that our objective is to use this description in real-life computations. From the computational viewpoint, when we try to save as much computer memory as possible, it is advantageous not to store easily recoverable information such as all the pairs $\langle S, v\rangle$ corresponding to the cases when S is inconsistent or when $v \in S$. Since we do not wan to waste computer memory on storing this non-informative information, we decided not to include it into our description of a measuring device.

In short, the resulting definition may be somewhat clumsy from the logical viewpoint, but, hopefully, more natural from the viewpoint of describing measuring devices.

Comment. It is worth mentioning that the idea to restrict deduction systems to only consistent sets S is not new: it has been proposed by D. Scott in [72] in his description of an *information system*. From this viewpoint, the only difference between our definition and the definition of an information system is that Scott still allows $S \vdash v$ for $v \in S$ while we do not. So, if $\vdash$ is our relation, then we can define the information system $\vdash'$ as follows:

$$S \vdash' v \leftrightarrow (S \vdash v) \vee v \in S.$$

Vice versa, if we have an information system with a deduction relation $\vdash'$, we can define the $v \notin S$-type relation $\vdash$ as follows:

$$S \vdash v \leftrightarrow (S \vdash' v) \,\&\, v \notin S.$$

Description in terms of existing mathematical structures: domains. We know that the information about subsets can be described in term of Girard domains. How does the additional conditional information look like in these terms?

For example, the statement "if v_1 and v_3 are both observed for some object, then v_2 will be observed as well" means that the class v_1v_3 – of all objects for which both v_1 and v_3 are observed – coincides with the class $v_1v_2v_3$ – of all objects for which all three outcomes v_1, v_2, and v_3 are observed.

In terms of the domain partial order $\sqsupseteq$, this equality can be described as follows. In the general case of a Girard domain, we have relations of the type $v_1v_2v_3 \sqsupseteq v_1v_3$ – meaning that when we observe v_1, v_2, and v_3 for the same object, we gain all the information about this object that could be gained from observing only v_1 and v_3 (and maybe some more information).

In the case of conditional statements, we have additional relations of the type $v_1v_3 \sqsupseteq v_1v_2v_3$ – meaning that when we observe v_1 and v_3 for the same object, we gain all the information about this object that could be gained from observing all three outcomes v_1, v_2, and v_3.

In general, in domain terms, a conditional statement $\langle S, v\rangle \in \mathcal{C}$ means, crudely speaking, that $S = S \cup v$.

Such additional relations bring us outside the class of Girard domains, to the class of general domains. Thus, we can say that while subsets of compatible outcomes correspond to *Girard domains*, situation when conditional statements are also available corresponds to *general domains*.

Comment. The fact that we started with a notion of a deduction system similar to Scott's information systems and ended up with a general domain is not accidental: information systems were explicitly defined as a logical representation of general domains.

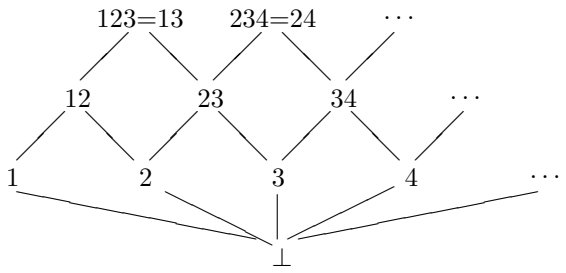
Let us illustrate this structure on the examples of typical measuring devices.

Example 1: interval uncertainty. For the simplest interval uncertainty, there exists a constant C such that v_i and v_j are compatible if and only if $|i - j| \le C$.

For $C = 1$, each point is only connected to its immediate neighbors, and no three outcomes are compatible. In principle, in cases when no three outcomes are compatible, we could have conditional statement like "if v_i is observed for some object, then v_j will eventually be observed for this object"; however, in the specific case $C = 1$, no such statements are true, so this case is described by the same Girard domain as in the subset case.

For $C = 2$, each element is compatible not only with its immediate neighbors, but also with immediate neighbors of immediate neighbors. Here, every triple $(i, i+1, i+2)$ is compatible; no quadruples are compatible. In this case, there are non-trivial conditional statements: namely, for every i, once we have observed v_{i-1} and v_{i+1}, we know that eventually, we will be able to observe v_i as well. Thus, in this case, $v_{i-1}v_iv_{i+1} = v_{i-1}v_{i+1}$.

Because of these conditional statements, instead of the original Girard domain, we have the following domain structure:



Example 2: counting. For counting, there are no conditional statements, so we still have a Girard domain.

Example 3: "yes"-"no" measurements. For the general case of "yes"-"no" measurements, there are no conditional statements, so we still have a Girard domain.

Example 4: a combination of several independent measuring instruments. When a measuring device consists of m independent measuring measurements, then the outcome x of the device is a tuple $(x_1, \ldots, x_m)$ of outcomes corresponding to individual measuring instruments.

Since the measurements are independent, a set of tuples $S \subseteq X_1 \times \ldots \times X_m$ implies a tuple $v = (v_1, \ldots, v_m)$ if and only if for each k, the corresponding measuring results $\pi_k(S)$ imply v_k.

In these terms, if we have m structures $\langle X_i, \mathcal{S}_i, \mathcal{C}_i\rangle$ corresponding to m different measuring instruments, then the measuring device can be described by a structure $\langle X, \mathcal{S}, \mathcal{C}\rangle$, where $X \stackrel{\rm def}{=} X_1 \times \ldots \times X_m$,

$$S \in \mathcal{S} \leftrightarrow \pi_1(S) \in \mathcal{S}_1 \,\&\, \ldots \,\&\, \pi_m(S) \in \mathcal{S}_m,$$

and

$$\langle S, v\rangle \in \mathcal{C} \leftrightarrow \langle \pi_1(S), \pi_1(v)\rangle \in \mathcal{C}_1 \,\&\, \ldots \,\&\, \langle\langle \pi_m(S), \pi_m(v)\rangle \in \mathcal{C}_m.$$

Thus defined structure will be called a *Cartesian product* of the structures $\langle X_1, \mathcal{S}_1, \mathcal{C}_1\rangle$, ..., $\langle X_m, \mathcal{S}_m, \mathcal{C}_m\rangle$.

Computational complexity of the domain representation of a measuring device: a general case. How many bits do we need to store the new information about the measuring device? We have already shown that to store the general simplicial complex information, we need at least $\approx 2^n$ bits. Since a simplicial complex is a particular case of the new representation – corresponding to the case when there are no valid conditional statements – we therefore need at least 2^n bits to represent the general domain information as well.

Let us show that 2^n bits are sufficient. If we store the information about $\mathcal{C}$ directly, then we may need more than 2^n bits: indeed, there are 2^n possible sets S, and n possible outcomes v, so overall, there are $2^n \cdot n$ pairs $\langle S, v\rangle$. Thus, if we describe $\mathcal{C}$ by storing the information about every such pair whether it belongs to $\mathcal{C}$ or not, we will need $2^n \cdot n$ bits, which is more than $O(2^n)$.

This pair-wise representation would be excessive, because due to superset and transitivity properties, once we know some pairs $\langle S, v\rangle$ from $\mathcal{C}$, we can automatically conclude that several other pairs also belong to $\mathcal{C}$.

Let us show that by using these properties, we can indeed reduce the required amount of storage to $O(2^n)$.

Specifically, instead of storing the sets $\mathcal{S}$ and $\mathcal{C}$ directly, we will store the set B of the corresponding n-dimensional *Boolean vectors*. A Boolean vector describes which of the n variables are true and which are not; in other words, it is a vector whose components are truth values. A natural way to describe a Boolean vector is by a sequence of the type $v_1^{\varepsilon_1} \dots v_n^{\varepsilon_n}$, where $\varepsilon_i \in \{-, +\}$, $v_i^+ \stackrel{\text{def}}{=} v_i$, and $v_i^- \stackrel{\text{def}}{=} \bar{v}_i$ – a negation of v_i. For example, for $n = 3$, an expression $v_1\bar{v}_2v_3$ means that v_1 and v_3 are true and v_2 is false.

Intuitively, $b \in B$ means that there exists an object for which all variables v_i that are true in b can occur in a measurement and all the variables v_i that are false in b cannot occur. For example, $v_1\bar{v}_2v_3 \in B$ would mean that there exists an object for which v_1 and v_3 are possible outcomes but the outcome v_2 will never occur.

We will construct B as follows: a Boolean vector b belongs to B if and only if the following two properties hold for the set S of all variables v_i that are true in b:

- the set S is *compatible*, i.e., $S \in \mathcal{S}$;
- the set S is *closed under deduction*, i.e., there are no statements of the type $\langle S, v\rangle$ in the set $\mathcal{C}$.

Both properties make perfect sense in view of our intended meaning of Boolean vectors $b \in B$ as describing the results of all possible measurement for some fixed object o:

- Since the true variables from b are exactly the outcomes that appear as a result of measuring o, the set S of all such true variables must be compatible. This justifies the first property.
- If $\langle S, v\rangle \in \mathcal{C}$ for some $v \notin S$, this means that for every object – including our object o – once we observed all the outcomes from the set S, we *must* also *observe* the outcome v. However, v does not belong to the set S of all the variables that are true in b; thus, v is false in b – meaning that the outcome v will *never be observed* for this object. This contradiction shows that no such pair $\langle S, v\rangle$ is possible – this justifies the second property.

Overall, there are 2 Boolean values, hence $2 \cdot 2 = 2^2$ Boolean vectors of length 2, ..., and $2 \cdot \ldots \cdot 2$ (n times) $= 2^n$ possible n-dimensional Boolean vectors. To

store the information about a general set B of Boolean vectors, it is sufficient to store, for each of 2^n possible Boolean vectors b, the information on whether $b \in B$. For each b, we need 1 bit to store this information, thus, overall, we need 2^n bits.

Let us show that storing the set B is sufficient in the sense that once we know B, we can uniquely reconstruct $\mathcal{S}$ and $\mathcal{C}$. Indeed, as one can easily prove, $S \in \mathcal{S}$ if and only if there exists a vector b in which all variables from S are true. Similarly, $\langle S, v\rangle \in \mathcal{C}$ if and only if for for every vector b in which all variables from S are true, v is also true. Thus, 2^n bits are sufficient to store the general domain information as well.

Comment. This Boolean representation has a direct meaning in logical terms. Namely, storing $\mathcal{S}$ and $\mathcal{C}$ would mean that we consider all possible implications that form a logical *theory*. By using Boolean vectors, we consider all possible *models* of this theory, i.e., all possible Boolean vectors that satisfy all the statements (conditional and unconditional) from this theory.

Computational complexity of the domain representation of a measuring device: case of interval uncertainty. We already know that for interval uncertainty, we need $O(n)$ bits to store the information $\mathcal{S}$ about compatible subsets S. Let us show that we need $O(n)$ bits to store the information $\mathcal{C}$ about conditional statements $\langle S, v\rangle$.

Indeed, for interval uncertainty, every outcome i is only compatible with neighbors that are C-close to i. Thus, for every conditional statement $\langle S, v\rangle$ in which $i \in S$, the compatible set S is a subset of the set $\{i-C, i-C+1, \ldots, i-1, i, i+1, \ldots, i+C\}$ consisting of $2C+1$ elements, and v must be an element of this set (in a conditional statement, v is compatible with every outcomes from S hence with i as well). For each i from 1 to n, there are finitely many ($\le 2^{2C+1}$) such subsets S and finitely many ($\le 2C+1$) such elements v. So, we need finitely many bits to store all the conditional statements in which $i \in S$, and thus, $O(n)$ bits to store the information about all conditional statements.

Thus, overall, to store the information about both $\mathcal{S}$ and $\mathcal{C}$, we need $O(n) + O(n) = O(n)$ bits.

Computational complexity of the simplicial complex representation of a measuring device: case of convex multi-D uncertainty. For every pair $\langle S, v\rangle \in \mathcal{C}$, the set $S \cup \{v\}$ is compatible. We have already shown that in the m-dimensional convex case, every compatible set has no more than $m+1$ outcomes, and thus, there are no more than $O(n^{m+1})$ such sets. Once we know the set $S \cup \{v\}$ consisting of $\le m+1$ elements, there are $\le m+1$ different ways of separating this set into S and v – i.e., $\le m+1$ pairs $\langle S, v\rangle$ that can possibly belong to $\mathcal{C}$. For each of $O(n^{m+1})$ compatible sets, there are $\le m+1$ such pairs – thus, the overall number of such pairs does not exceed $(m+1) \cdot O(n^{m+1}) = O(n^{m+1})$.

In order to get a complete description of $\mathcal{C}$, it is therefore sufficient to know, for each of these $O(n^{m+1})$ pairs, whether this particular pair belongs to $\mathcal{C}$. Thus, to store all the information about $\mathcal{C}$, we need $O(n^{m+1})$ bits.

We already know that we need $O(n^{m+1})$ bits to store the information about $\mathcal{S}$. Thus, overall, to store all the information about $\mathcal{S}$ and $\mathcal{C}$, we need $O(n^{m+1}) + O(n^{m+1}) = O(n^{m+1})$ bits.

Computational complexity of the domain representation of a measuring device: general case of localized uncertainty. In the case of localized uncertainty, each outcome has no more than C neighbors. To store the information about these neighbors, as we have mentioned in the graph case, we need $O(n \cdot \log(n))$ bits. To store the information about the simplicial complex structure, we need $O(n)$ more bits.

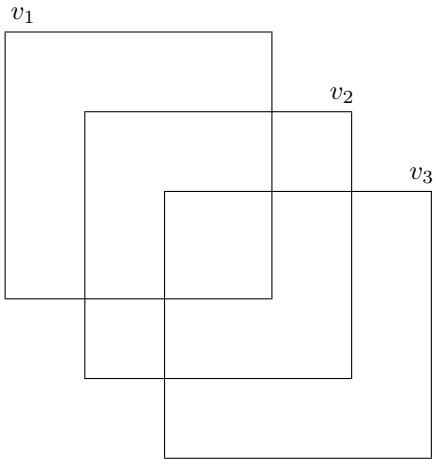
How many more bits do we need to store the conditional statements from the set $\mathcal{C}$? For each outcome x, for every conditional statement $\langle S, v\rangle \in \mathcal{C}$ for which $x \in S$, the set S is a subset of the set of x's neighbors, and v is one of x's neighbors. For each of n outcomes x, there are finitely many ($\leq 2^C$) such subsets S and finitely many ($\leq C$) such neighbors v, thus, overall, there are finitely many ($\leq 2^C \cdot C$) pairs $\langle S, v\rangle \in \mathcal{C}$ for which $x \in S$. To store the information on which of these pairs do belong to $\mathcal{C}$ and which do not, we need finitely many bits.

Thus, in the localized case, we need $O(n)$ bits to store the complete information about $\mathcal{C}$. Overall, we need $O(n \cdot \log(n)) + O(n) + O(n) = O(n \cdot \log(n))$ bits – the same amount as in the graph case.

2.6 Fifth Step: Disjunctive Conditional Statements about the Possible Outcomes

Addition of conditional statements does not always lead to a complete description of a measuring device. In the previous subsection, we have shown that the addition of conditional statements sometimes provides more information about a measuring device than the (unconditional) information on which sets of physically possible outcomes are compatible and which are not – and thus provide a *more* complete description of a measuring device. However, as we will see, this description is still not always fully complete.

To show this incompleteness, let us consider yet another pair of measuring devices each of which has three possible outcomes v_1, v_2, and v_3. The first device is the same interval-related device as in the previous subsection, in which v_1 corresponds to interval $S_1 = [-1, 3]$, v_2 corresponds to $S_2 = [0, 4]$, and v_3 corresponds to $S_3 = [1, 5]$. In the second device, the uncertainty corresponding to each outcome is described by a box (2-D interval):



In both devices, the three sets S_i corresponding to the outcomes v_i have a common point, hence all three outcomes are compatible. In both devices, there is one conditional statement: if, for some object o, we have observed v_1 and v_3, then we will eventually observe v_2 for this same object. So, from the viewpoint of the conditional statements description (= domain description), the two measuring devices are described by exactly the same structure $\langle X, \mathcal{S}, \mathcal{C}\rangle$, where $X = \{v_1, v_2, v_3\}$,

$$\mathcal{S} = \{\{v_1\}, \{v_2\}, \{v_3\}, \{v_1, v_2\}, \{v_1, v_3\}, \{v_2, v_3\}, \{v_1, v_2, v_3\}\},$$

and $\mathcal{C} = \{\langle\{v_1, v_3\}, v_2\rangle\}$. However, from the physical viewpoint, there is a difference between these devices, a difference that is not captured by the above common description.

Namely, for the second (box-related) measuring device, there are values that belong to S_2 but do not belong neither to S_1 nor to S_3. This means that it is possible to have an object for which v_2 is a possible outcome, but for which neither v_1 nor v_3 are possible outcomes.

In contrast, for the first (interval-related) measuring device, v_2 is contained in the union of v_1 and v_3. Thus, when for some object, a measurement produces v_2, we are sure that further repeated measurement will eventually produce either the outcome v_1 or the outcome v_3 – if not for this particular device, at least for one of the devices characterized by the same uncertainty.

What we do we need to add to the conditional statements description to capture the missing information about a measuring device? We have shown that there is a difference between the measuring devices that is not captured by the previously considered conditional statements. How can we describe such a difference in general terms?

In the first (interval-related) measuring device, for every object, if v_2 is a possible outcome, then either v_1 or v_3 is also a possible outcome. In the second

measuring device, there exist objects for which v_2 is a possible outcome but neither v_1 not v_3 ever occur.

To capture the above difference, we must also formalize *disjunctive conditional* statements of this type, i.e., statements of the type

"If for some object, v_2 is a possible outcomes, then for this same object, either v_1 or v_3 is also a possible outcome."

In general, we may have a set S of outcomes, a set U of outcomes, and need a statement of the following type:

"If for some object, all outcomes from a set $S \subseteq X$ have been observed, then one of the outcomes $v \in U$ will also be eventually observed for this object."

In the particular case when U is a one-element set $U = \{v\}$, we get conditional statements of the type considered in the previous subsection.

The existence of a full theory makes the set of all true disjunctive conditional statements algorithmically listable. Similarly to the previous subsection, we can argue that the existence of a full theory leads to the possibility to tell, for each potential disjunctive conditional statement, whether this statement is true for this particular measuring device.

In other words, for every two sets $S, U \subseteq X$, we can tell whether the above conditional statement is true. Thus, we can conclude that the set of all true statement of this type is algorithmically listable.

Family of true disjunctive conditional statements: natural properties. Each of the statements and the entire family of such statements must satisfy the following natural properties.

First, in describing each individual disjunctive conditional statement, it is sufficient to consider only sets U for which the following two properties hold:

- $S \cap U = \emptyset$ (otherwise,the conditional statement is trivially true), and
- all elements of U are compatible with s, i.e., for all $v \in U$, $S \cup \{v\} \in \mathcal{S}$ (otherwise, by definition of $\mathcal{S}$, once S has been observed, we cannot observe v for the same object).

To summarize: once a family $\mathcal{S}$ is given, we can say that a pair $\langle S, U\rangle$, where $S, U \in \mathcal{S}$, is a *possible disjunctive condition* if $U \cap S = \emptyset$ and $S \cup \{v\} \in \mathcal{S}$ for all $v \in U$.

Second, if S implies U, and we have observed $S' \supseteq S$, this means that we have, in particular, observed S, so S' also implies U.

Also, if S implies that one of the outcomes from U will be observed, and $U' \supseteq U$, this means that one of the outcomes from U' will be thus observed – so S also implies U'. We can combine this property with the previous one and conclude that S' implies U'.

Third,

- if S implies $U \cup \{v\}$ – meaning that whenever all outcomes from S are observed for some object, then either v will be eventually observed, or one of the outcomes from U will be observed, and
- if for some other two sets S' and U', $S' \cup \{v\}$ implies U' – meaning that once we observed v and all outcomes from S', then we will eventually observe one of the outcomes from U',

then we can conclude that once we observe all the outcomes from S and from S', we will eventually observe either one of the outcomes from U, or (if S leads to v) one of the outcomes from U', i.e., we an conclude that $S \cup S'$ implies $U \cup U'$.

Thus, we arrive at the following conclusion.

Conclusion: algorithmically listable family of disjunctive conditional statements. We conclude that at the current (fifth) step of describing a measuring device, we can describe it as a triple $\langle X, \mathcal{S}, \mathcal{C}\rangle$, where:

- X is a finite set (of all physically possible outcomes),
- $\mathcal{S} \subseteq 2^X$ is a subset of 2^X that satisfies the following two properties:
  - *subset property* (if $S \in \mathcal{S}$ and $S' \subseteq S$, then $S' \in \mathcal{S}$), and
  - the *singleton property* ($\mathcal{S}$ contains all singleton sets),

  and
- $\mathcal{C}$ is a family of pairs $\langle S, U\rangle$, $S, U \subseteq X$, that satisfies the following three properties:
  - Every pair $\langle S, U\rangle \in \mathcal{C}$ is a *possible disjunctive condition*, i.e., $U \cap S = \emptyset$ and $S \cup \{v\} \in \mathcal{S}$ for all $v \in U$.
  - *superset property*: If $\langle S, U\rangle \in \mathcal{C}$, $S' \supseteq S$, $U' \supseteq U$, and $\langle S', U'\rangle$ is a possible disjunctive condition, then $\langle S', U'\rangle \in \mathcal{C}$; and
  - *transitivity* (*cut*, *resolution*): if $\langle S, U \cup \{v\}\rangle \in \mathcal{C}$, $\langle S' \cup \{v\}, U'\rangle \in \mathcal{C}$, and $\langle S \cup S', U \cup U'\rangle$ is a possible disjunctive condition, then $\langle S \cup S', U \cup U'\rangle \in \mathcal{C}$.

Description in terms of existing mathematical structures: sequent calculus. From the mathematical viewpoint, the set $\mathcal{C}$ of pairs $\langle S, U\rangle$ has the same properties as the *sequent calculus* $\vdash$, i.e., the relation $S \vdash U$ meaning that once all the statement from S are true, one of the statements from U must also be true; see, e.g., [5, 23, 49, 66]. In sequent calculus, what we called a subset property corresponds to the weakening rules, and what we called transitivity is called a *cut* or, in more modern times, a *resolution rule*.

Comment. It is worth mentioning that the resolution rule is the basis for most modern systems of automatic reasoning, and also for logic programming and its use in Artificial Intelligence.

Comment. Similarly to the case of deduction relation, what we describe in slightly different from the traditional sequent calculus:

- first, we only allow consistent sets S;
- second, we do not consider cases $S \cap U \neq \emptyset$ when the disjunctive condition is always true.

These cases are easy to add to our description, because, e.g., if S is inconsistent, then it implies everything ($S \vdash U$ for all U). Thus, from the mathematical viewpoint, our definition describes exactly the same structures as the traditional one.

Description in terms of existing mathematical structures: Boolean vectors. In describing computational complexity of the conditional statements description, we have already seen that it is beneficial to describe conditional statements in terms of a set B of possible Boolean vectors b. In other words, it is beneficial, instead of considering all possible implications that form a logical *theory*, to consider all possible *models* of this theory, i.e., all possible Boolean vectors that satisfy all the statements (conditional and unconditional) from this theory.

In view of this advantage, let us therefore use the same idea in our more general case, when we also have disjunctive conditional statements.

As before, $b \in B$ will mean that there exists an object for which all variables v_i that are true in b can occur in a measurement and all the variables v_i that are false in b cannot occur.

Thus, a Boolean vector b belongs to B if and only if the following two properties hold for the set S of all variables v_i that are true in b:

- the set S is *compatible*, i.e., $S \in \mathcal{S}$;
- the set S is *closed under deduction*, i.e., there are no statements of the type $\langle S, U\rangle$ in the set $\mathcal{C}$.

Once we know B, we can uniquely reconstruct $\mathcal{S}$ and $\mathcal{C}$. Indeed, as one can easily prove, $S \in \mathcal{S}$ if and only if there exists a vector b in which all variables from S are true. Similarly, $\langle S, U\rangle \in \mathcal{C}$ if and only if for for every vector b in which all variables from S are true, one of the variables from U is also true.

Thus, Boolean vectors indeed provide a possible description of disjunctive conditional statements.

Example. Let us show how this description looks like on the example of the above two measuring devices. For the first (interval-related) measuring device, the outcome v_1 is possible for $x \in S_1 = [-1, 3]$, the outcome v_2 is possible for $x \in S_2 = [0, 4]$, and the outcome v_3 is possible for $x \in S_3 = [1, 5]$. Thus, we have the following Boolean vectors:

- when x is between -1 and 0, only the outcome v_1 is possible, which corresponds to the Boolean vector $v_1 \overline{v}_2 \overline{v}_3$;

- when x is between 0 and 1, the outcomes v_1 and v_2 are possible, and v_3 is not possible, which corresponds to the Boolean vector $v_1 v_2 \bar{v}_3$;
- when x is between 1 and 3, all three outcomes are possible, which corresponds to the Boolean vector $v_1 v_2 v_3$;
- when x is between 3 and 4, the outcomes v_2 and v_3 are possible, and v_1 is not possible, which corresponds to the Boolean vector $\bar{v}_1 v_2 v_3$;
- finally, when x is between 4 and 5, only the outcome v_3 is possible, which corresponds to the Boolean vector $\bar{v}_1 \bar{v}_2 v_3$.

In other words, this measuring device can be described by the set

$$B_1 = \{v_1 \bar{v}_2 \bar{v}_3, v_1 v_2 \bar{v}_3, v_1 v_2 v_3, \bar{v}_1 v_2 v_3, \bar{v}_1 \bar{v}_2 v_3\}.$$

For the second (box-related) measuring device, it is also possible that v_2 occurs but v_1 and v_3 do not occur, so the corresponding set B_2 of Boolean vectors contains not only all five vectors from the above set B_1, but also an additional Boolean vector $\bar{v}_1 v_2 \bar{v}_3$.

Comment. A Boolean vector is uniquely determined by the set $S \subseteq X$ of its true-valued components. Thus, describing a set of Boolean vectors is equivalent to describing a family $\mathcal{F}$ of subsets $S \subseteq X$. In this sense, this description is similar to the Girard domain description. There are two differences, however, between the subset version of Boolean vector description and the Girard domain description:

- First, Girard domain had a subset property while our description does not have a subset property. Indeed, it is quite possible that, say, a set $\{v_1, v_2\}$ belongs to $\mathcal{F}$ – meaning that the corresponding Boolean vector $v_1 v_2 \bar{v}_3$ is possible, but its subset $\{v_2\}$ does not belong to $\mathcal{F}$ – meaning that the corresponding Boolean vector $\bar{v}_1 v_2 \bar{v}_3$ is not possible.
- Second, the property that every outcome from the set X is physically possible, the property that was described, in Girard domains, by the singleton set property, is now described by a different property: for every $v \in X$, there exists a set $S \in \mathcal{F}$ for which $v \in S$.

Description in terms of existing mathematical structures: Boolean algebra. The description in terms of Boolean vectors can be reformulated in more algebraic terms. Indeed, the fact that, say, a vector $v_1 \bar{v}_2 v_3$ belongs to the set B of possible vectors means that there exists an object for which v_1 and v_3 are possible but v_2 is not. In other words, it means that the intersection of the sets S_1 (corresponding to v_1), S_3 (corresponding to v_3), and a complement $-S_2$ of the set S_2 (corresponding to v_2) is non-empty: $S_1 \cap (-S_2) \cap S_3 \neq \emptyset$.

We already know that the need for intersections leads to domains. In general, to describe Boolean vectors in set-theoretic terms, we need not only intersection,

but also the *complement* operation. There is a mathematical notion that is related to intersection and complement in the same way as domains are related to intersection only: the notion of a *Boolean algebra*; see, e.g., [37, 39, 56].

In the finite case, the Boolean algebra can be defined similarly to how we defined domains: an *algebra of sets* can be defined as a family of sets that is closed under intersection and complement, and a *Boolean algebra* can be defined as a partially ordered set (with an additional operation "complement") that is isomorphic to an algebra of sets.

Of course, just like for domains, there are also direct definitions of Boolean algebras, definitions that are more useful in proving theorems, but for our motivational purposes, the above indirect definition works fine.

In Boolean algebra, since we have intersection and complement, we can also define union $A \cup B$ as $-((-A) \cap (-B))$. From the viewpoint of a measuring device, a union $S_1 \cup S_2$ means that we know that we got either v_1 or v_2 as an outcome.

This "or" statement does not correspond to a direct measurement by a measuring device: in such a direct measurement, we always know which outcome we got. However, as we mentioned in the Introduction, we are also interested in functions $f(x_1, \ldots, x_n)$ that, based on the outcome $\widetilde{x}_i$ of one (or several) measuring device, try to predict the outcome $\widetilde{y}$ of measuring by another (difficult-to-use) measuring device. In some such situations, we may not be able to predict the outcome $\widetilde{y}$, but we may be able to provide some information about $\widetilde{y}$ – meaning that some outcomes $v_1, \ldots, v_m$ are possible. In such situations, the prediction is $v_1 \vee \ldots \vee v_m$ – i.e., the union $S_1 \cup \ldots \cup S_m$ of the corresponding sets.

For a single measurement, we can either predict the outcome $\widetilde{y}$ exactly, or we can predict a union (disjunction) of possible outcomes. Since, as we have mentioned, we can gain additional information by performing repeated measurements, it is reasonable to see what we can predict about *repeated* measurement of y. The results of each repeated measurement can be described by which outcomes will occur and which do not, i.e., by a conjunction (intersection) of the type $v_1 \,\&\, \bar{v}_2 \,\&\, v_3$. If we cannot make an exact prediction, we can therefore predict that one of such sequences will be observed. In other words, a general prediction about a repeated measurement is a disjunction (union) of possible conjunctions (intersections).

In mathematical logic, such disjunctions of conjunctions have a special name – Disjunctive Normal Forms (DNF, for short). It is known that an arbitrary propositional formula – i.e., a formula obtained from n propositional ("yes"-"no") variables $v_1, \ldots, v_n$ by using operations $\&$ ("and"), $\vee$ ("or"), and $\bar{v}_i$ ("not") – can be represented in the equivalent DNF form. Thus, possible predictions about a repeated measurement are arbitrary propositional formulas formed from the variables $v_1, \ldots, v_n$. By definition, the class of all propositional formulas is closed under $\&$, $\vee$, and negation.

In set theoretic terms, $\&$ correspond to the intersection, $\vee$ corresponds to the union, and negation corresponds to the complement. Thus, the class of all the sets corresponding to different predictions is closed under union, intersection,

and complement – i.e., it forms an *algebra of sets* isomorphic to the Boolean algebra.

Once we know which Boolean vectors b belongs to the set B of possible vectors, we can uniquely describe this algebra: a Boolean vector is in B if and only if the corresponding intersection is non-empty, these intersections do not intersect with each other, and an arbitrary set from the algebra of sets – corresponding to an arbitrary propositional formula F – is a union of the intersections corresponding to all Boolean vectors for which F is true.

Example. For the first (interval-related) measuring device, we have 5 basic Boolean vectors, so we have $2^5 = 32$ different elements in the Boolean algebra, elements corresponding to different subsets of the set B_1. For example, the subset $\{v_1\bar{v}_2\bar{v}_3, v_1v_2\bar{v}_3\}$ corresponds to a propositional formula $v_1\bar{v}_2\bar{v}_3 \vee v_1v_2\bar{v}_3$ (this formula can be simplified into $v_1\bar{v}_3$).

For the second (box-related) measuring device, we have 6 Boolean vectors, so we have $2^6 = 64$ different elements in the corresponding Boolean algebra.

Description in terms of existing mathematical structures: domains. The same information can be described in domain terms: namely, from the Boolean algebra, we only pick up propositional formulas that do not contain negation.

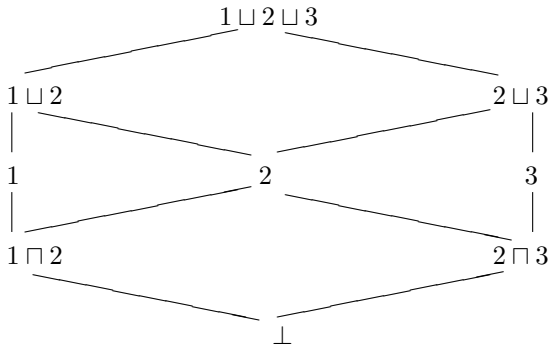
Every such formulas can be represented in a DNF form, as a disjunction of conjunctions of variables, such as $v_1v_3 \vee v_2v_4$.

Comment about notations. We have already mentioned that the conjunction v_iv_j of outcomes v_i and v_j corresponds to the intersection $S_i \cap S_j$ of the corresponding sets and, in terms of the domain order $x \sqsupseteq y \leftrightarrow x \subseteq y$, a *join* (least upper bound) $x \sqcup y$. Similarly, the disjunction $v_i \vee v_j$ of outcomes v_i and v_j corresponds to the union $S_i \cup S_j$ of the corresponding sets and, in terms of the domain order $x \sqsupseteq y \leftrightarrow x \subseteq y$, a *meet* (greatest lower bound) $x \sqcap y$.

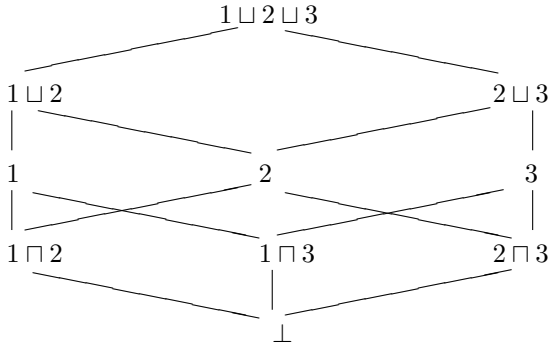
In these domain notations, the above formula takes the form $v_1v_3 \sqcap v_2v_4$, or, equivalently, $(v_1 \sqcup v_3) \sqcap (v_2 \sqcup v_4)$.

Let us show that this description is equivalent to the original one in the sense that one description enables us to uniquely reconstruct the other one. Indeed, once we know a Boolean algebra, we can always retain only positive elements and get a domain. Vice versa, once we know, for every two "positive" elements, whether they contain each other, we will be able to uniquely reconstruct the set B of possible Boolean vectors and thus, the original Boolean algebra. Indeed, e.g., $v_1\bar{v}_2v_3\bar{v}_4 \notin B$ means that $S_1 \cap (-S_2) \cap S_3 \cap (-S_4) = \emptyset$ which is equivalent to $S_1 \cap S_3 \subseteq S_2 \cup S_4$, i.e., to the domain ordering relation $v_1v_3 \sqsupseteq v_2 \vee v_4$.

Example. For the first (interval-related) measuring device, we get the following domain structure:



For the second (box-related) measuring device, we have an additional element $1 \vee 3$ in this powerset domain. Indeed, for this device, if we know that either v_1 or v_3 is the outcome, we still get some non-trivial information about the measured quantity. In contrast, for the interval-related measuring device, $v_1 \vee v_3$ is always true and thus, knowing that this is true would gain us no information at all:



Comment. Until now, we considered a domain in which elements are compatible sets of measurement outcomes, such as $v_1 v_3$. In the new domain, possible elements are disjunctions of such sets, such as $v_1 v_3 \vee v_2 v_4$. In other words, elements of the new domain are *subsets* of the old domain. The set of all subsets is usually called a *powerset*, so the elements of the new domain form a powerset over the set of all the elements of the old domain. It is therefore reasonable to call the new domain a *powerset domain*.

Our notion of a powerset domain is somewhat similar to the notion of a "power domain" that has been extensively studied in domain theory. However, the notion of a power domain was studied mainly in situations when, once we know the domain, we can uniquely construct the power domain. There are

several different definition of a power domain, but once we fix a definition, the power domain is uniquely determined.

In contrast, the main reason for introducing the new feature is our example showing, in effect, that for domains corresponding to measuring devices, we cannot uniquely reconstruct a powerset domain from the original domain. Thus, the same domain can have several different powerset domains.

Is this a final description of validated uncertainty? This is the fifth step in our description of a measuring device. In previous steps, we proposed reasonable descriptions of a measuring device – only to discover later that each of these descriptions does not capture all the information about this measuring device. So, a natural question is: is this fifth step description final or some further improvements are possible?

We will argue that as far as validated uncertainty is concerned, this description is final. Indeed, from the viewpoint of validated uncertainty, for each given object, we do not worry about the probabilities of different outcomes (i.e., about how often these outcomes appear), we are only interested in knowing which outcomes will appear and which outcomes will never appear. From this viewpoint, for each object, all we are interested in is a Boolean vector that tells us, for each of n physically possible outcomes v of a measuring device, whether this outcome can occur for this object or not.

So, for each object, the (validated part of the) behavior of the measuring device is fully characterized by the corresponding Boolean vector. Overall, to fully characterize a measuring device, we thus need to know all Boolean vectors corresponding to all possible objects – and this is exactly our description B.

Example 1: interval uncertainty. For the simplest interval uncertainty, there exists a constant C such that v_i and v_j are compatible if and only if $|i-j| \le C$. All such structures are similar, so it is sufficient to consider the case when $v_1 = 1$ and $h = 1$; in this case, $v_k = k$.

For $C = 1$, i.e., for $\Delta = 0.5$, the outcome v_1 corresponds to the interval $S_1 = (-\infty, 1.5]$, the outcome v_n corresponds to the interval $S_n = [n - 0.5, \infty)$, and every intermediate outcome v_k corresponds to the interval $S_k = [k-0.5, k+0.5]$. In this case, we have the following Boolean vectors:

- when x is smaller than 0.5, only v_1 is possible, so we have $v_1 \bar{v}_2 \dots \bar{v}_n$;
- when x is equal to 0.5, both v_1 and v_2 are possible, so we have $v_1 v_2 \bar{v}_3 \dots \bar{v}_n$;
- when x is between 0.5 and 1.5, only v_2 is possible, so we have $\bar{v}_1 v_2 \bar{v}_3 \dots \bar{v}_n$;

  ...
- when x is equal to $k - 0.5$, both v_{k-1} and v_k are possible, so we have $\bar{v}_1 \dots \bar{v}_{k-2} v_{k-1} v_k \bar{v}_{k+1} \dots \bar{v}_n$;
- when x is between $k - 0.5$ and $k + 0.5$, only v_k is possible, so we have $\bar{v}_1 \dots \bar{v}_{k-1} v_k \bar{v}_{k+1} \dots \bar{v}_n$;

...

- finally, when x is larger than $n - 0.5$, only v_n is possible, i.e., we have $\bar{v}_1 \ldots \bar{v}_{n-1} v_n$.

In this case, the set B of all Boolean vectors consists of the following $2n - 1$ vectors:

- n vectors $\bar{v}_1 \ldots \bar{v}_{k-1} v_k \bar{v}_{k+1} \ldots \bar{v}_n$ $(1 \le k \le n)$ in which only one component is true, and
- $n - 1$ vectors $\bar{v}_1 \ldots \bar{v}_{k-2} v_{k-1} v_k \bar{v}_{k+1} \ldots \bar{v}_n$ $(2 \le k \le n)$ in which two neighboring components are true-valued.

For $C = 2$, i.e., for $\Delta = 1$, the outcome v_1 corresponds to the interval $S_1 = (-\infty, 2]$, the outcome v_n corresponds to the interval $S_n = [n - 1, \infty)$, and every intermediate outcome v_k corresponds to the interval $S_k = [k - 1, k + 1]$. In this case, we have the following Boolean vectors:

- when x is smaller than 1, only v_1 is possible, so we have $v_1 \bar{v}_2 \ldots \bar{v}_n$;
- when x is between 1 and 2, both v_1 and v_2 are possible, so we have $v_1 v_2 \bar{v}_3 \ldots \bar{v}_n$;
- when x is equal to 2, three outcomes are possible: v_1, v_2, and v_3, so we have $v_1 v_2 v_3 \bar{v}_4 \ldots \bar{v}_n$;

  ...
- when x is equal to k, three outcomes are possible: v_{k-1}, v_k, and v_{k+1}, so we have $\bar{v}_1 \ldots \bar{v}_{k-2} v_{k-1} v_k v_{k+1} \bar{v}_{k+2} \ldots \bar{v}_n$;
- when x is between k and $k + 1$, only v_k and v_{k+1} are possible, so we have $\bar{v}_1 \ldots \bar{v}_{k-1} v_k v_{k+1} \bar{v}_{k+2} \ldots \bar{v}_n$;

  ...
- finally, when x is larger than n, only v_n is possible, i.e., we have $\bar{v}_1 \ldots \bar{v}_{n-1} v_n$.

In this case, the set B of all Boolean vectors consists of the following $2n - 1$ vectors:

- 2 vectors $v_1 \bar{v}_2 \ldots \bar{v}_n$ and $\bar{v}_1 \ldots \bar{v}_{n-1} v_n$ in which only one component is true-valued;
- $n - 1$ vectors $\bar{v}_1 \ldots \bar{v}_{k-1} v_k v_{k+1} \bar{v}_{k+2} \ldots \bar{v}_n$ $(1 \le k \le n - 1)$ in which two neighboring components are true-valued, and
- $n - 2$ vectors $\bar{v}_1 \ldots \bar{v}_{k-2} v_{k-1} v_k v_{k+1} \bar{v}_{k+2} \ldots \bar{v}_n$ $(2 \le k \le n - 1)$ in which three neighboring components are true-valued.

Example 2: counting. For counting, no two outcomes are compatible, so we have exactly n possible Boolean vectors $\bar{v}_1 \ldots \bar{v}_{k-1} v_k \bar{v}_{k+1} \ldots \bar{v}_n$ $(1 \le k \le n)$ in each of which exactly one component is true.

Example 3: "yes"-"no" measurements. In "yes"-"no" measurements, we measure, e.g., a quantity x with an accuracy Δ in order to decide, based on the measurement result $\widetilde{x}$, whether $x \le x_0$. In this case, we have the following possibilities:

- When $x \le x_0 - 2\Delta$, we thus have $\widetilde{x} \le x_0 - \Delta$ and hence, knowing that $|\widetilde{x} - x| \le \Delta$, we can definitely conclude that $x \le x_0$. In this case, the only possible outcome is 0, so the corresponding Boolean vector is $0\bar{U}\bar{1}$.
- When $x_0 - 2\Delta \le x \le x_0$,
  - we can have $\widetilde{x} \le x_0 - \Delta$ – in which case we can still still conclude that the answer is 0,
  - and we can also have $\widetilde{x} > x_0 - \Delta$ – in which case, based on $\widetilde{x}$, we cannot decide whether $x \le x_0$, so the answer is U.

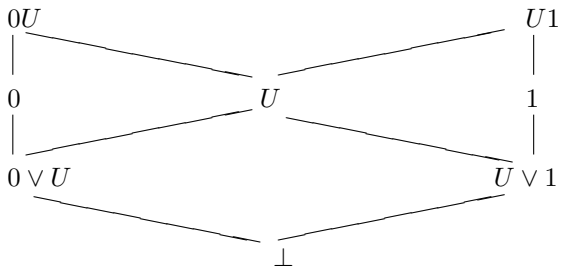
  So, possible outcomes are 0 and U, so the Boolean vector is $0U\bar{1}$.
- When $x_0 < x \le x + 2\Delta$,
  - we can have $\widetilde{x} > x_0 + \Delta$ – in which case we can conclude that $x > x_0$ and thus, that the answer is 1,
  - and we can also have $\widetilde{x} \le x_0 + \Delta$ – in which case, based on $\widetilde{x}$, we cannot decide whether $x \le x_0$, so the answer is U.

  So, possible outcomes are 1 and U, so the Boolean vector is $\bar{0}U1$.
- Finally, when $x > x_0 + 2\Delta$, then we have $\widetilde{x} > x_0 + \Delta$, hence, knowing that $|\widetilde{x} - x| \le \Delta$, we can definitely conclude that $x > x_0$. In this case, the only possible outcome is 1, so the corresponding Boolean vector is $\bar{0}\bar{U}1$.

Thus, we get the following set of possible Boolean vectors:

$$B = \{0\bar{U}\bar{1}, 0U\bar{1}, \bar{0}U1, \bar{0}\bar{U}1\}.$$

The resulting powerset domain looks as follows:



Example 4: a combination of several independent measuring instruments. When a measuring device consists of m independent measuring measurements, then the outcome x of the device is a tuple $(v_1, \ldots, v_m)$ of outcomes corresponding to individual measuring instruments.

Let B_i be the set of Boolean vectors $b : X_i \to \{0, 1\}$ corresponding to i-th individual instrument. For each object, we have m Boolean vectors $b^{(i)} \in B_i$ $(1 \le i \le m)$ corresponding to m measuring instruments. A vector $b^{(i)}$ describes possible outcomes of applying i-th measuring instrument to this object: an outcome $v_i \in X_i$ is possible if and only if $b^{(i)}(v_i) = 1$ (= "true").

Since the measurements are independent, for each object, a combination $(v_1, \ldots, v_m)$ is possible if and only if each v_i is possible. Thus, for n Boolean vectors $b^{(i)}$, the resulting Boolean vector is as follows:

$$b(v_1, \ldots, v_m) = b^{(1)}(v_1) \,\&\, \ldots \,\&\, b^{(m)}(v_m).$$

We will call such a vector a *tensor product* of the Boolean vectors $b^{(i)}$ and denote it by $b^{(1)} \otimes \ldots \otimes b^{(m)}$.

So, if we have m structures $\langle X_i, B_i \rangle$ corresponding to m different measuring instruments, then the measuring device can be described by a structure $\langle X, B \rangle$, where $X = X_1 \times \ldots \times \ldots X_m$ and

$$B = \{b^{(1)} \otimes \ldots \otimes b^{(m)} \,|\, b^{(1)} \in B_1 \,\&\, \ldots \,\&\, b^{(m)} \in B_m\}.$$

Computational complexity of the Boolean representation of a measuring device: a general case. On this stage, an information about a measuring device with n possible outcomes is represented as a set of n-dimensional Boolean vectors. We already know, from the analysis of the computational complexity of the previous step, that this representation requires 2^n bits, and in general, it is necessary to use all these bits.

Computational complexity of the Boolean representation of a measuring device: case of interval uncertainty. For interval uncertainty, every outcome i is only compatible with neighbors that are C-close to i. For every Boolean vector b, we can select the first true-valued component as i. Thus, for every Boolean vector in which i-th is the first true-valued component (i.e., in which i-th outcome is possible and none of the outcomes $1, \ldots, i-1$ are possible), the only other true-valued components must form a subset of the set $\{i, i+1, \ldots, i+C\}$ consisting of $C+1$ elements – all more distant components are false.

Since in every Boolean vector, only finitely many components are true-valued, it makes sense, instead of storing the entire Boolean vector, to only store the list S of its true-valued components. To save space even further, it makes sense to store, for each i, only the differences between the corresponding values and i, so that, e.g., the set $\{i, i+1, i+2\}$ is store as $\{0, 1, 2\}$.

For each i from 1 to n, there are finitely many $(\le 2^{2C+1})$ such subsets S that starts with i, and each of these subsets contains no more than $C+1$ elements

of size $\leq C$. So, we need finitely many bits to store all the Boolean vectors in which i is the first true value, and thus, $O(n)$ bits to store the information about all the Boolean vectors.

Computational complexity of the Boolean representation of a measuring device: case of convex multi-D uncertainty. For every Boolean vector $b \in B$, the set S of all its true-valued components is compatible. We have already shown that in the m-dimensional convex case, every compatible set has no more than $m+1$ outcomes, and thus, there are no more than $O(n^{m+1})$ such sets.

To get a complete description of B, it is therefore sufficient to know, for each of these $O(n^{m+1})$ sets, whether the corresponding Boolean vector belongs to B. Thus, to store all the information about B, we need $O(n^{m+1})$ bits.

Computational complexity of the domain representation of a measuring device: general case of localized uncertainty. In the case of localized uncertainty, each outcome has no more than C neighbors. To store the information about these neighbors, as we have mentioned in the graph case, we need $O(n \cdot \log(n))$ bits.

How many more bits do we need to store the information about all the Boolean vectors B? For each outcome x, and for every Boolean vector that is true on x, the set S of all positive components is a subset of the set of x's neighbors. For each of n outcomes x, there are finitely many ($\leq 2^C$) such subsets S. To store the information on which of these subsets correspond to Boolean vectors from B, we need finitely many bits.

Thus, in the localized case, we need $O(n)$ bits to store the complete information about B. Overall, we need $O(n \cdot \log(n)) + O(n) = O(n \cdot \log(n))$ bits – the same amount as in the previous cases.

2.7 Summary

Descriptions of a measuring device corresponding to all 5 steps can be summarized in the following table:

Step #	Description	Corresponding mathematical structure	Corresponding class of domains
1	set of possible outcomes X	set	–
2	set of compatible pairs $\langle X, \sim \rangle$	graph	binary Girard domains
3	family of compatible sets of outcomes $\langle X, \mathcal{S} \rangle$	simplicial complex	Girard domains
4	conditional statements $\langle X, \mathcal{S}, \mathcal{C} \rangle$	deduction relation $\vdash$	general domains
5	set of possible Boolean vectors $\langle X, B \rangle$	sequent calculus, Boolean algebra	powerset domains

Computational complexity of different representations can be summarized in the following table:

Step #	Description	General case	Simple interval uncertainty	m-D convex uncertainty	Localized uncertainty
1	possible outcomes	$O(n)$	$O(n)$	$O(n)$	$O(n)$
2	compatible pairs	$O(n^2)$	$O(n)$	$O(n^2)$	$O(n \cdot \log(n))$
3	compatible sets of outcomes	$O(2^n)$	$O(n)$	$O(n^{m+1})$	$O(n \cdot \log(n))$
4	conditional statements	$O(2^n)$	$O(n)$	$O(n^{m+1})$	$O(n \cdot \log(n))$
5	Boolean vectors	$O(2^n)$	$O(n)$	$O(n^{m+1})$	$O(n \cdot \log(n))$

2.8 Measuring Device: A Final Description

A measuring device can be described as a pair $\langle X, B \rangle$, where X is a finite set with $|X|$ elements and $B \subset \{0,1\}^X$ is a set of $|X|$-dimensional Boolean vectors.

Elements of the set X are physically possible outcomes. A Boolean vector b belongs to the set B if and only there exists an object for which repeated measurements lead to all outcomes $v \in X$ for which $b(v) = 1$ and for which the outcomes v with $b(v) = 0$ never occur.

Comment. On each of the five steps, we argued that the existence of a full theory should enable us to produce the information corresponding to the corresponding step:

- on Step 1, we argue that the existence of a full theory describing individual measurement results enable us to list all physically possible outcomes;
- on Step 2, we argue that the existence of a full theory describing pairs measurement results enables us to list all compatible pairs,
- on Step 3, we argue that the existence of a full theory describing sequences of repeated measurement results enables us to list all compatible subsets,

etc.

On each step, we require more and more of a full theory. If we do have a full theory corresponding to our final Step 5, then it is reasonable to use the above final description. However, it is possible that while we have, e.g., a full theory describing all pairs of measurement results, we do not yet have a full theory describing all physically possible sequences of measurement results. In this case, the only description of a measuring device that we can produce is its description as a graph.

With this possibility in mind, in the following text, we will consider not only a measuring device as a pairs $\langle X, B\rangle$ corresponding to the final Step 5, but we will also consider descriptions corresponding to all previous steps of our description.

3 Physical Quantities: A General Description

General idea. How can we describe a general physical quantity? The value of the quantity is obtained from measurements, so it is natural to describe a quantity in terms of measurements.

In the ideal situation when there are no measurement errors, the measurement result is exactly equal to the value of the measured quantity. In this case, the set of physically possible outcomes of the measuring device is exactly the set of values of the physical quantity.

In real life, measurements are rarely 100% accurate. In most real-life situations, the measurement result is only approximately equal to the actual value of the measured quantity. As a result, the set of physically possible outcomes of a measuring device provides us with only an approximate description of the measured quantity.

To get a more accurate description of a physical quantity, we must have, in addition to the original measuring instrument, a second, more accurate measuring instrument. When we apply both measuring instruments to the same object, we get a more accurate representation of the measured quantity. From our viewpoint, we can view both instruments as forming a "virtual" measuring device that provides a more accurate description of the desired physical quantity.

Comment. In some cases, it may be difficult to find a single measuring instrument that will provide more accurate measurements for the entire scale. Instead, we may have different measuring instruments for different parts of the scale. For

example, we may need different sensors for accurately measuring temperatures close to an absolute zero and temperatures close the body temperature. In such situations, instead of adding a single measuring instrument, we may need to add several measuring instruments. In this case, the virtual measuring device consists of more than two measuring instruments.

Since measuring instruments are rarely perfect, the combined measuring device is also providing us only with an approximate value of the measured quantity – and thus, its set of physically possible outcomes provides us only with an approximate description of the measured quantity. To get a yet better description, we must use a third even more accurate measuring instrument, i.e., we must add a third measuring instrument to our virtual measuring device.

If the quality of this third device is not sufficient, we may need a fourth device, etc. At each stage of this process of more and more accurate measurement, we get an approximate description of the measured quantity. The actual value of the measured quantity can be viewed as a "limit" of these approximate description – limit corresponding to the case when the measurement error tends to 0.

From the general idea to a formal description. To describe a physical quantity, we need a sequence of measuring devices, each of which is obtained from the previous one by adding one or more measuring instruments. In other words, we have measuring instruments $I^{(1)}$, $I^{(2)}$, etc., corresponding to more and more accurate measurements, and we form measuring devices $I^{(1)}$, $I^{(1)}I^{(2)}$, $I^{(1)}I^{(2)}I^{(3)}$, ..., corresponding to more and more accurate measurements.

Usually, there are several different measuring instruments at each accuracy level, so when we combine them in different ways, we do not get a single *sequence*, we get a *family* of measuring devices. For example, if we have two measuring instruments $I_1^{(2)}$ and $I_2^{(2)}$ on the 2nd accuracy level, then, instead of just two possible measuring devices $I^{(1)}$ and $I^{(1)}I^{(2)}$, we can form several devices: $I^{(1)}$, $I^{(1)}I_1^{(2)}$, $I^{(1)}I_2^{(2)}$, and $I^{(1)}I_1^{(2)}I_2^{(2)}$.

Comment. When we simply combine several measuring instruments into a single measuring device, then, from the physical viewpoint, the result does not depend on the order in which we combine these instruments. From this viewpoint, we will not distinguish between, e.g., $I^{(1)}I_1^{(2)}I_2^{(2)}$ and $I^{(1)}I_2^{(2)}I_1^{(2)}$.

We already know how to describe a single measuring device: as a set of physically possible outcomes X equipped with an additional information describing possible outcomes of repeated measurements. To formalize our idea of a physical quantity, we must also be able to describe, in these terms, what it means that one measuring device I (with set X of physically possible outcomes) is obtained from another one I' (with set X' of physically possible outcomes) by adding some extra measuring instrument(s) E' (we will denote the corresponding set of possible outcomes also by E').

Let us follow all 5 steps of our description of a measuring instrument and see how can we describe this relation on each of these steps.

Set of possible outcomes: the notion of a projection. When we combine a measuring instrument I' with an extra measuring instrument E' into a single measuring device I, then each outcome v of the resulting measuring device I is a pair $v = (v', e')$, where $v' \in X'$ is the outcome of the first measuring instrument and $e' \in E'$ is the outcome of the second measuring instrument.

The set V of physically possible outcomes of the combined measuring device is thus a set of such pairs, i.e., a subset of the Cartesian product $X \subseteq X' \times E'$.

Since we know that each outcome v' from the set X' is physically possible, this means that for every $v' \in X'$ there exists a value $e' \in E'$ for which $(v', e') \in X$. Similarly, since we know that each outcome e' from the set E' is physically possible, this means that for every $e' \in E'$ there exists a value $v' \in X'$ for which $(v', e') \in X$.

Once we know the outcome $v = (v', e') \in X$ of the combined measuring device I, we can reconstruct the outcome v' of the first measuring instrument I' by simply taking the first component v' and ignoring the second result e'. So, we have a natural function $(v', e') \to v'$ from X to X'.

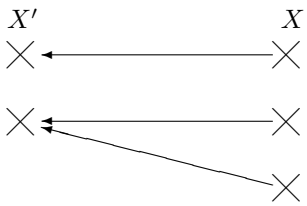
This function can be naturally described in geometric terms. In standard measurements, when v' and e' are real numbers, the pair (v', e') is naturally represented by a point (v', e') on the plane, with x-coordinate v' and y-coordinate e'. In these terms, the function $(v', e') \to v'$ that maps every point on the plane into its x-coordinate is a *projection* on the x-axis. Projections are usually denoted by π. Because of this geometric interpretation, we will use the same term *projection* (and use the π notation) to describe the general mapping of the type $(v', e') \to v'$.

Comment. In domain theory, another term (borrowed from category theory) is also often used to describe a mapping that simply erases ("forgets") one or more components of the original information – a *forgetful functor*.

How can we describe a projection in general terms? In other words, if we only know the sets X and X', what are the conditions that a mapping $f : X \to X'$ should satisfy so that it is possible to describe a set X as a set of pairs $X \subseteq X' \times X''$ for which $f(v', e') = v'$?

Not every mapping $f : X \to X'$ can be thus represented: e.g., we know that all elements of X' must appear in the image $f(X) = \{f(x) \,|\, x \in X\}$ of this mapping, i.e., that $f(X) = X'$. It turns out that once a function $f : X \to X'$ satisfies this property, we can always find a set E' and a 1-1 equivalence between X and an appropriate subset of $X' \times E'$ for which $f(X)$ is simply a projection $(v', e') \to v'$. Indeed, we can take $E' = X$ and represent each $v \in X$ as a pair $(f(v), v) \in X' \times X''$; in this representation, f is clearly a projection.

In view of this observation, if we want to guarantee that a function $f : X \to X'$ is a projection, then the only property that we need to satisfy is the property $f(X) = X'$.

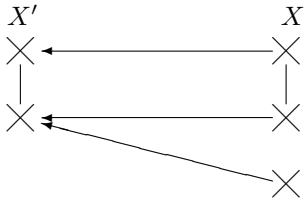


Thus, we arrive at the following definition:

Definition. A mapping $f : X \to X'$ is called a *projection* if $f(X) = X'$, i.e., if it is an "onto" mapping (surjection).

Pairs of compatible outcomes: the notion of a projection. For pairs of compatible outcomes, if the outcomes $v_1 = (v_1', e_1')$ and $v_2 = (v_2', e_2')$ are compatible, this means that there is an object for which both outcomes v_1 and v_2 are possible. For this same object, the first measuring instrument I' produces outcomes $v_1' = f(v_1)$ and $v_2' = f(v_2)$. By definition of a compatibility relation, this means that v_1' and v_2' are compatible. Thus, if $v_1 \sim v_2$, then $f(v_1) \sim' f(v_2)$.

Vice versa, if v_1' and v_2' are compatible, this means that there exists an object for which repeated measurements with the measuring instrument I' produce outcomes v_1' and v_2'. Let e_1' and e_2' denote the corresponding results of applying the extra measuring instrument E' to the same object. Thus, for the combined measuring device I, we get the outcomes $v_1 = (v_1', e_1')$ and $v_2 = (v_2', e_2')$. Since we get these outcomes by measuring the same object, we have $v_1 \sim v_2$. Thus, whenever $v_1' \sim' v_2'$, there exists values $v_1 \sim v_2$ for which $f(v_1) = v_1'$ and $f(v_2) = v_2'$.



So, we arrive at the following definition:

Definition. Let $M = \langle X, \sim \rangle$ and $M' = \langle X', \sim' \rangle$ be graphs. A mapping $f : X \to X'$ is called a *projection* of M onto M' if the following three properties are satisfied:

- $f(X) = X'$ – i.e., f is a projection in the set theoretic sense;
- if $v_1 \sim v_2$, then $f(v_1) \sim' f(v_2)$ – i.e., f preserves $\sim$, and
- if $v_1' \sim' v_2'$, there exists $v_1 \sim v_2$ for which $f(v_1) = v_1'$ and $f(v_2) = v_2'$.

Subsets of compatible outcomes: the notion of a projection. For the case when a measuring device is represented by subsets of compatible outcomes, we arrive at a similar definition:

Definition. Let $M = \langle X, \mathcal{S}\rangle$ and $M' = \langle X', \mathcal{S}'\rangle$ be simplicial complexes. A mapping $f : X \to X'$ is called a *projection* of M onto M' if the following three properties are satisfied:

- $f(X) = X'$ – i.e., f is a projection in the set theoretic sense;
- if $S \in \mathcal{S}$, then $f(S) \in \mathcal{S}'$ – i.e., f preserves compatibility, and
- if $S' \in \mathcal{S}'$, there exists a set $S \in \mathcal{S}$ for which $f(S) = S'$.

Comment. For every family of sets $\mathcal{S} \subset 2^X$ and for every function $f : X \to X'$, we can define $f(\mathcal{S})$ as

$$f(\mathcal{S}) \stackrel{\text{def}}{=} \{f(S) \,|\, S \in \mathcal{S}\}.$$

This notation enables us to simplify the above notion of a projection: a function $f : X \to X'$ is a projection if $f(X) = X'$ and $f(\mathcal{S}) = \mathcal{S}'$ – i.e., if it is a projection in set theoretic sense both as a function from X to X' and as a function from $\mathcal{S}$ to $\mathcal{S}'$.

Definition reformulated in domain terms. In domain terms, instead of the sets X and X', we have domains with, correspondingly, $\mathcal{S}$ as the set of all elements of the first domain and $\mathcal{S}'$ as the set of all elements of the second domain. Elements of the original set X (or X') are minimal elements (different from $\perp$) of the corresponding domain.

One of the properties of the projection is that f maps compatible sets into compatible sets, so f is a mapping from $\mathcal{S}$ to $\mathcal{S}'$. This mapping must satisfy the following three properties:

- be a projection in a set-theoretic sense,
- map minimal elements into minimal elements, and
- map union into union: if $a \sqcup b \in \mathcal{S}$, then $f(a \sqcup b) = f(a) \sqcup f(b)$.

Comment. It is worth mentioning that such mapping are continuous, stable, and linear in the sense of Girard domains; see, e.g., [78].

Vice versa, if the above three conditions are satisfied for a function $f : \mathcal{S} \to \mathcal{S}'$, then f is defined on singleton sets, and its values on singletons are also singletons. For every $v \in X$, the image of the singleton set $\{v\}$ is also a singleton set, i.e., a set consisting of only one element. Let us denote this single element by $g(v)$. Then, for every $v \in X$, we have $f(\{v\}) = \{g(v)\}$. Due to the

union property, for every non-singleton set $S = \{s_1, \ldots, s_m\} = \{s_1\} \cup \ldots \{s_m\}$, we have

$$f(S) = f(\{s_1\}) \cup \ldots \cup f(\{s_m\}) = \{g(s_1)\} \cup \ldots \cup \{g(s_m)\} =$$

$$\{g(s_1), \ldots, g(s_m)\} = g(S).$$

Thus, f indeed defines a projection.

Therefore, in Girard domain terms, we can define a projection as a mapping from $\mathcal{S}$ to $\mathcal{S}'$ that is a projection, maps minimal elements into minimal, and satisfies the union property.

General domains and Boolean vectors: the notion of a projection. As we have mentioned, the case of general domains can be described in terms of Boolean vectors, so it is sufficient to describe a projection for Boolean vectors.

In a Boolean vector description, a measuring device is represented by a set B of Boolean vectors. This set consists of vectors $b \in B$ corresponding to different objects. For every object, its possible outcomes are represented by a Boolean vector b in which $b(v) = 1$ (= "true") if v is possible (and will eventually occur) for this object, and $b(v) = 0$ (= "false") if v is impossible (and will never occur) for this object.

From this viewpoint, if $b \in B$ is a possible Boolean vector for the combined measuring device $I = I'E'$, then for the corresponding object, possible outcomes $v = (v', e')$ are exactly those for which $b(v) = 1$. Thus, for this object, an outcome v' is possible for the measuring instrument I' if $b(v', e') = 1$ for some $e' \in E'$, i.e., if $b(v) = 1$ for some v for which $f(v) = v'$. In other words, the corresponding Boolean vector b' for I' has the following form: $b'(v') = 1$ if and only if $b(v) = 1$ for some v for which $f(v) = v'$.

In algebraic terms, we can say that

$$b'(v') = \max_{v: f(v) = v'} b(v).$$

The set B' should therefore be equal to the image of the set B under this transformation. Thus, we arrive at the following definition:

Definition. By a *Boolean structure*, we mean a pair $\langle X, B\rangle$, where X is a set, and B is a set of Boolean vectors, i.e., of mappings $X \to \{0, 1\}$. Let $\langle X, B\rangle$ and $\langle X', B'\rangle$ be two Boolean structures, and let $f : X \to X'$ be a projection (in set-theoretic sense). For every Boolean vector $b : X \to \{0, 1\}$, we define $f(b) : X' \to \{0, 1\}$ as follows:

$$(f(b))(v') = \max_{v: f(v) = v'} b(v).$$

We say that a mapping $f : X \to X'$ is a *projection* of the corresponding Boolean structures if f is a projection in set-theoretic sense (i.e., $f(X) = X'$) and $f(B) = B'$.

Comment. If we represent each Boolean vector b by the set of its true-valued components $S = \{v \mid b(v) = 1\}$, then the above definition becomes even simpler: if the Boolean vector b is represented by a set S, then $f(b)$ is represented by the set $f(S)$. So, in this representation, the definition of a projection is similar to the definition of a projection for Girard domains.

Comment. A Boolean vector puts into correspondence, to every element v, a value 1 or 0 (true or false). Boolean vectors can be viewed as a particular case of *fuzzy sets* (see, e.g., [43, 59]) in which we put into correspondence, to every v, a number from the interval $[0, 1]$. Here, 0 corresponds to "false", 1 corresponds to "true", and intermediate values correspond to intermediate degrees of uncertainty.

From this viewpoint, it is worth mentioning that the above definition of how to extend a function $f : X \to X'$ to Boolean vectors is, in effect, a particular case of the more general definition of how to extend a function to fuzzy sets. This definition was originally proposed by L. Zadeh, the founder of fuzzy logic, and it is usually called *Zadeh's extension principle.*

The coincidence of these two extensions is not accidental: it is known that fuzzy sets can be described in terms of embedded families of sets (in particular, intervals), and that this representation explains the extension principle; see, e.g., [58].

Open question. Boolean vectors can be naturally represented in domain terms: namely, in terms of powerset domains. We have seen that for Girard domains, the notion of a projection can be naturally reformulated in domain terms. It is desirable to find a similar domain reformulation of this more general Boolean-based notion of a projection.

The family of all measuring devices measuring a given physical quantity: a description. We have already mentioned that, because of inevitable measurement errors, if we want to get an accurate description of a physical quantity, it is not sufficient to have a single measuring device measuring this quantity. Instead, we need to consider a (potentially infinite) family of all measuring devices that measure this quantity.

For this family to be sufficient, for each measuring device I', this family must contain a more accurate (more informative) device I, i.e., a device that includes, in addition to all the sensors that are already in I', some additional (more accurate) sensors. In other words, on the family of all measuring devices, there is a natural relation "a measuring device I is obtained from another one I' by adding extra measuring instrument(s)", a relation that we described above as the existence of a projection $\pi_{I,I'} : I \to I'$.

The notion "more informative" is, as we have mentioned, the fundamental notion of domain theory. Thus, it is reasonable to describe this relation "I is more informative than I'" by domain-style notations $I \sqsupseteq I'$ or, equivalently, $I' \sqsubseteq I$.

What are the natural properties of this relation? First, let us show that this relation is transitive. Indeed, if I' is a projection of I, this means that I' is obtained from I by ignoring the results of some of the measuring instruments (sensors) comprising the measuring device I. Similarly, if I'' is a projection of I', this means that I'' is obtained from I' by ignoring the results of some of the measuring instruments (sensors) comprising the measuring device I'. It is therefore natural to conclude that I'' can be obtained from I by ignoring the results of some of the measuring instruments – i.e., that I'' is a projection of I. In other words, the relation $\sqsubseteq$ is transitive.

In terms of projection, transitivity means that if we have a function $\pi_{I,I'}$ that projects I onto I', and we have a function $\pi_{I',I''}$ that projects I' onto I'', then we have a function $\pi_{I,I''}$ that projects I onto I''. Let us show that the resulting projection is actually a composition of the projections $\pi_{I,I'}$ and $\pi_{I',I''}$.

Indeed, since I'' is a projection of I', each outcome of I' can be represented as $v' = (v'', e'')$, where v'' is the corresponding outcome of $I''$and e'' are the extra values that we ignore when we consider I'. Similarly, each outcome of I can be represented as (v', e'), where e' are the corresponding extra outcomes. Since $v' = (v'', e'')$, every outcome of I has the form $((v'', e), e')$, i.e., the form (v'', e'', v'). The projection $\pi_{I,I'}$ maps (v', e') into v', i.e., (v'', e'', v') into (v'', e''). The projection $\pi_{I',I''}$ maps (v'', e'') into v''. The projection $\pi_{I,I''}$ maps (v'', e'', e') directly into v''. We can see that for every outcome $v = (v'', e'', e') \in X$, we have $\pi_{I',I''}(\pi_{I,I'})(v) = \pi_{I,I''}(v)$, i.e., that the projection $\pi_{I,I''}$ is indeed a composition of the projections $\pi_{I,I'}$ and $\pi_{I',I''}$: $\pi_{I,I''} = \pi_{I,I'} \circ \pi_{I',I''}$.

Comment. It is easy to see that for all above definitions, the composition of projections is indeed a projection.

Another property of the relation $\sqsubseteq$ between measuring devices is that it is *directed*. Indeed, for every two such measuring devices I' and I'', we can always combine them into a single measuring device I. Thus, for every I' and I'', there exists a measuring device I for which $I' \sqsubseteq I$ and $I'' \sqsubseteq I$. So, we arrive at the following definition:

Definition. By a *family of measuring devices*, we mean a pair $\langle \mathcal{I}, \mathcal{P} \rangle$, where:

- $\mathcal{I}$ is a denumerable set of measuring devices I, I', ..., and
- $\mathcal{P}$ is a set of projections $\pi_{I,I'}$ corresponding to some pairs of measuring devices $I, I' \in \mathcal{I}$,

that satisfies the following two properties:

- *transitivity:* if $\pi_{I,I'} \in \mathcal{P}$ and $\pi_{I',I''} \in \mathcal{P}$, then $\mathcal{P}$ contains a projection $\pi_{I,I''} = \pi_{I,I'} \circ \pi_{I',I''}$;
- *directedness:* for every $I', I'' \in \mathcal{I}$, there exists a measuring device $I \in \mathcal{I}$ for which $\pi_{I,I'} \in \mathcal{P}$ and $\pi_{I,I''} \in \mathcal{P}$.

Physical quantity as a projective limit of measuring devices. The whole purpose of the family of (more and more accurate) measuring devices is that, if we want to know the actual value of the physical quantity, we should apply these devices one after another. After each additional measurement, we will get a more and more accurate description of the actual value. The actual value can thus be viewed as a "limit" of these approximate results. How can we describe this limit?

When we apply the first measuring device I' to the quantity, we get some outcome v' from the set X' of physically possible outcomes of this measuring device. If we then take the second measuring device I that extends I' (in the sense that $\mathcal{P}$ contains a projection $\pi_{I,I'} : I \to I'$), then we get an outcome $v \in X$. By definition of a projection, the original result v' is a projection of the new result v on I', i.e., $\pi_{I,I'}(v) = v'$.

Thus, after the first two measurements, the measurement results can be described as a mapping that maps each measuring device from the family $\{I, I'\}$ into one of its possible outcomes in such a way that $\pi_{I,I'}(v) = v'$.

To get a more accurate description of the measured value, we can involve more measuring devices from our family $\mathcal{I}$. At any given moment of time, we can only involve finitely many devices, but eventually, we can involve each of them.

Thus, we can describe the actual value of the measured quantity as a mapping a that maps every measuring device $I \in \mathcal{I}$ with the set of possible outcomes X_I into a value $a_I \in X_I$ in such a way that $\pi_{I,I'}(a_I) = a_{I'}$ for all projections $\pi_{I,I'} \in \mathcal{P}$.

The set of all such mappings is called a *projective limit* of the family $\langle \mathcal{I}, \mathcal{P} \rangle$. Thus, we can define the range of a physical quantity as such a limit.

Definition. Let $\langle \mathcal{I}, \mathcal{P} \rangle$ be a family of measuring devices. By a *value of the corresponding measured physical quantity*, we mean a mapping a that map every measuring device $I \in \mathcal{I}$ with the set of possible outcomes X_I into a value $a_I \in X_I$ in such a way that $\pi_{I,I'}(a_I) = a_{I'}$ for all projections $\pi_{I,I'} \in \mathcal{P}$. The set X of all such values will be called the *range* of the corresponding quantity, or, alternatively, the *projective limit* of the family of measuring devices and denoted by

$$X = \lim_{\leftarrow} X_I.$$

Example. Let us illustrate this definition on the example of a simple case case when $\mathcal{I}$ is a sequence of (more and more accurate) measuring devices $I_1, I_2, \ldots, I_n, \ldots$, in which each device is a refinement of the previous one, i.e., $\pi_{I_{i+1},I_i} \in \mathcal{P}$ for every i. This structure can be described as follows:

$$I_1 \overset{\pi_{I_2,I_1}}{\longleftarrow} I_2 \overset{\pi_{I_3,I_2}}{\longleftarrow} \ldots \overset{\pi_{I_{n-1},I_n}}{\longleftarrow} I_{n-1} \overset{\pi_{I_n,I_{n-1}}}{\longleftarrow} I_n \overset{\pi_{I_{n+1},I_n}}{\longleftarrow} I_{n+1} \overset{\pi_{I_{n+2},I_{n+1}}}{\longleftarrow} \ldots$$

Each element of the projective limit can be described as a mapping that maps each measuring device I_n into an element v_n from the set X_n of possible out-

comes of this device in such as way that for every n, $\pi_{I_{n+1},I_n}(v_{n+1}) = v_n$:

$$v_1 \overset{\pi_{I_2,I_1}}{\longleftarrow} v_2 \overset{\pi_{I_3,I_2}}{\longleftarrow} \ldots \overset{\pi_{I_{n-1},I_n}}{\longleftarrow} v_{n-1} \overset{\pi_{I_n,I_{n-1}}}{\longleftarrow} v_n \overset{\pi_{I_{n+1},I_n}}{\longleftarrow} v_{n+1} \overset{\pi_{I_{n+2},I_{n+1}}}{\longleftarrow} \ldots$$

In the general case, we arrive at the following definition:

Comment. In the definition of the range of a physical quantity, do we really need *all* the mappings a? For example, if we measure a 1-dimensional quantity with better and better accuracy, then (as we will see later) we get a family of narrower and narrower intervals that contain the actual real number – the value of the measured quantity.

Are all real numbers indeed physically possible? We do not know. It may be that the actual (unknown) value is always a rational number. It may also be that the actual value is always an irrational number. Our point is that we will never know, and thus, we can as well consider all possible real numbers.

Indeed, at any given moment of time, we only know the results of finitely many measurements. For example, in measuring a real-valued quantity, after a finite number of measurements, we only have an approximate description of the actual value, with some accuracy ε. Within this accuracy, no matter what measurement value $\widetilde{x}$ we got, it will always be possible that the actual (unknown) value $x \in [\widetilde{x} - \varepsilon, \widetilde{x} + \varepsilon]$ is rational, and it is always possible that this number is irrational.

We can assume that all the values are rational, but, no matter how many experiments we make, we will never be able to prove or disprove this assumption. In other words, from the physical viewpoint, this assumption can be neither verified nor falsified and thus, it has no relation with observations and measurements in the physical world. In short, such as assumption has as much physical meaning as a statement that there are invisible witches that can fly above us all the time, but are not detectable by any sensor.

In general, at any given moment of time, all we have is a result $v_I \in X_I$ for some measuring device I. Thus, whatever restrictions we make on the set of all possible sequences, there is no physical way for testing these restrictions – so we can as well assume (as we did) that all (appropriate) mappings a are possible.

Within this definition, the fact that every outcome from X_I is physically possible is now a theorem. Let us show that this definition is in good accordance with our understanding of X_I as the set of all physically possible domains. Indeed, we have the following result:

Proposition. Let $\langle \mathcal{I}, \mathcal{P} \rangle$ be a family of measuring devices, and let X be the set of all possible values of the corresponding measured physical quantity. Then, for every measuring device I and for every $v \in X_I$, there exists a value $a \in X$ for which $a_I = v$.

Proof. Let us denote I by I_0. Since the family $\mathcal{I}$ is denumerable, we can enumerate all its remaining elements into a sequence $\mathcal{I} - \{I_0\} = \{I_1, I_2, \ldots, I_n, \ldots\}$ Let us sequentially define the values $a_i = a_{I_i} \in X_i$, $i = 0, 1, \ldots$ of the desired mapping a. At each stage of this process, we will pick a marked device I_i and define the value of a_i for this device I_i and for several (finitely many) less informative devices I_j in such a way that for every l, n for which a_l and a_n are already defined and a projection exists, we have $\pi_{l,n}(a_l) = a_n$.

On the first stage, the definition is easy: we define $a_0 = v$. Let us now assume that we have already finished some stage, and we have thus defined the value a_i for some marked device I_i and the values a_j for several devices $I_j \sqsubseteq I_i$. Let k be the first integer for which the value a_k is not yet defined.

If I_k is less informative than the marked device I_i, i.e., if there is a projection $\pi_{i,k}$, then we can simply define $a_k \stackrel{\text{def}}{=} \pi_{i,k}(a_i)$ and keep the same device I_i marked. Due to transitivity, it is easy to see that after this definition, we still keep the projection property $\pi_{l,n}(a_l) = a_n$.

If I_k is not less informative than I_i, then one possibility is that I_k is *more* informative than I_i, i.e., that the function $\pi_{k,i}$ belongs to $\mathcal{P}$.

Since the function $\pi_{k,i} : X_k \to X_i$ is a projection and $a_i \in X_i$, there exists a value $a \in X_k$ for which $\pi_{k,i}(a) = a_i$. We will select this value a as a_k. We can then define a_k as $\pi_{m,k}(a_m)$ and select I_k as the new marked device. Due to transitivity, this addition also keeps the projection property $\pi_{l,n}(a_l) = a_n$.

Finally, we must consider the case when I_k is neither less informative nor more informative than I_i. In this case, we can use the fact that the family of a measuring devices is, by definition, a directed set. This means, in particular, that there exists a measuring device I_m that is at least as informative as both I_i and I_k, i.e., for which $\mathcal{P}$ contains both projections $\pi_{m,i}$ and $\pi_{m,k}$. In this case, we first add $I_m \sqsupseteq I_i$ and then $I_k \sqsubseteq I_m$. We know that each of these additions preserves the projection property.

Let us prove that eventually, we will thus define a_i for every measuring device I_i. Indeed, at each stage of this algorithm, we define the value a_k corresponding to the smallest integer k for which a_k was previously undefined. Thus, at each stage, we increase the smallest undefined value k at least by 1. Thus, eventually, we will indeed define all the values – and thus, define a mapping a for which $a_0 = v$. The proposition is proven.

Different sequences of measurement results may correspond to the same value of the measured quantity. So far, we have described an actual value of the measured quantity as a sequence (or, more generally, family) of measurement results corresponding to more and more accurate measuring devices.

What we have not yet taken into consideration is the fact that the same actual value of a physical quantity can lead to different sequences of measurement results – indeed, due to measurement uncertainty, even in the first measurement, we can get different outcomes if we repeatedly measure the same object. Thus, to complete the description of a physical quantity, we must also describe when

two sequences (families) correspond to the same actual value of the measured quantity.

The specifics of such a description depend on how we define a measuring device.

Case of graphs. Let us first consider the case when we describe each measurement device by a graph $\langle X, \sim\rangle$. In this case, if we have two mappings a and b corresponding to the same object, then, for every measuring device I, the corresponding measurement results must be compatible: $a(I) \sim_I b(I)$. It is therefore reasonable to say that the two mappings are *compatible* if $a(I) \sim_I b(I)$ for all I. It is also natural to say that mappings a and b are *indistinguishable* because, no matter how many measurements we make, we will never be able to conclude that one of them is different from another, there is always a possibility that these two families of measurement results represent the same actual value of the physical quantity.

Thus, we arrive at the following definition:

Definition. Let $\langle \mathcal{I}, \mathcal{P}\rangle$ be a family of measuring devices $I = \langle X_I, \sim_I\rangle$. By a *physical quantity*, we mean a pair consisting of this family of measuring devices and a graph $\langle X, \sim\rangle$, where:

- X is the projective limit of the sets X_I, and
- $a \sim b$ if and only if $a_I \sim_I b_I$ for all $I \in \mathcal{I}$.

Comment. This definition coincides with the definition of a binary Girard domain as a projective limit of Girard domain. So, from the viewpoint of *describing* the real world, it is sufficient to consider this limit domain as a description of the set of all possible values of the physical quantity. This is what we did in our previous papers [17, 18].

However, in this paper, our main interest is not in a general *description* of a physical quantity, but rather in what we can *compute* from the results of measuring this quantity. Every time we talk about measurement results, we talk about the outcomes of a specific measuring device $I \in \mathcal{I}$. Thus, in this paper, in addition to the limit domain, we also keep the corresponding family of measuring devices – i.e., the approximating domains.

In domain terms, this additional information is similar to the additional information used in the description of *effective* domains.

Within this definition, the fact that $\sim_I$ describes exactly compatible pairs is now a theorem. Let us show that this definition is in good accordance with our understanding of the relation $\sim_I$ as the relation describing which pairs of outcomes are compatible. Indeed, we have the following result:

Proposition. Let $\langle \mathcal{I}, \mathcal{P} \rangle$ be a family of measuring devices $I = \langle X_I, \sim_I \rangle$, and let $\langle X, \sim \rangle$ be the projective limit of this family. Then, for every measuring device I and for every two values $v, v' \in X_I$, the following two conditions are equivalent to each other:

- $v \sim_I v'$;
- there exist values $a, a' \in X$ for which $a \sim a'$, $a_I = v$, and $a'_I = v'$.

Proof. By definition of of the limit relation $\sim$, if $a \sim a'$, then $a_I \sim_I a'_I$, so the existence of such $a, a' \in X$ indeed implies that $v \sim_I v'$.

To complete the proof, we must therefore show that if $v \sim_I v'$, then there indeed exist the corresponding values $a, a' \in X$. We can do that by using a construction similar to how we proved that every value $v \in X_I$ is physically possible: namely, we enumerate the measuring devices into a sequence $I_0 = I$, I_1, I_2, ..., and then, on each stage, pick a marked device I_i and define a_i and a'_i for this device and for finitely many less informative devices.

Similarly, we start with $a_0 = v$ and $a'_0 = v'$. To extend our definition to a less informative device, we use projection. The possibility to extend to a more informative device follows from the fact that, by definition of a family of measuring instruments, when $I_i \sqsubseteq I_k$, the compatibility relation $\sim_i$ is a projection of the new relation $\sim_k$ – in the sense that whenever $a_i \sim_i a'_i$, there always exist values $a_k \sim_k a'_k$ for which $\pi_{k,i}(a_k) = a_i$ and $\pi_{k,i}(a'_k) = a'_i$.

Case of simplicial complexes. Let us now consider the case when we describe each measurement device by a simplicial complex $\langle X, \mathcal{S} \rangle$, where $S \subseteq 2^X$ has subset and singleton properties. In this case, if we have several mappings $S = \{a, b, \ldots, c\}$ corresponding to the same object, then, for every measuring device I, the corresponding measurement results must be compatible: $S(I) \stackrel{\text{def}}{=} \{a(I) \,|\, a \in S\} = \{a(I), b(I), \ldots, c(I)\} \in \mathcal{S}_I$. It is therefore reasonable to say that a finite set S of mappings is *compatible* (or *indistinguishable*) if $S(I) \in \mathcal{S}_I$ for all I. Thus, we arrive at the following definitions:

Definition. By a *simplicial complex*, we understand a pair $\langle X, \mathcal{S} \rangle$, where X is a set, and $\mathcal{S}$ is a family of finite subsets of X with the following two properties:

- *subset property*: if $S \in \mathcal{S}$ and $S' \subseteq S$, then $S' \in \mathcal{S}$;
- *singleton property*: for every $x \in X$, then $\{x\} \in \mathcal{S}$.

Comment. The difference from the previous definition of the simplicial complex (in the section in which we described measuring devices) is that previously, we were interested in the case when X is the set of all physically possible outcomes of a measuring device – and thus, a finite set. In our case, X is the set of all possible values of the measured quantity and thus, this set X can be infinite.

Definition. Let $\langle \mathcal{I}, \mathcal{P} \rangle$ be a family of measuring devices $I = \langle X_I, \mathcal{S}_I \rangle$. By a *physical quantity*, we mean a pair consisting of this family of measuring devices and a simplicial complex $\langle X, \mathcal{S} \rangle$, where:

- X is the projective limit of the sets X_I, and
- $S \in \mathcal{S}$ if and only if $S(I) \in \mathcal{S}_I$ for all $I \in \mathcal{I}$.

Comment. This definition coincides with the definition of a Girard domain as a projective limit of finite Girard domains.

Within this definition, the fact that $\mathcal{S}_I$ describes exactly compatible subsets is now a theorem. The above definition is in good accordance with our understanding of the class $\mathcal{S}_I$ as the class of all compatible sets of outcomes. Indeed, we have the following result:

Proposition. Let $\langle \mathcal{I}, \mathcal{P} \rangle$ be a family of measuring devices $I = \langle X_I, \mathcal{S}_I \rangle$, and let $\langle X, \mathcal{S} \rangle$ be the projective limit of this family. Then, for every measuring device I and for every set $S_I \subseteq X_I$, the following two conditions are equivalent to each other:

- S_I is compatible in I, i.e., $S_I \in \mathcal{S}_I$;
- there exists a finite set $S \subseteq X$ for which $S \in \mathcal{S}$ and $S(I) = S_I$.

Comment. The proof of this result is similar to the proof for pairs.

Cases of conditional statements and Boolean vectors. To explain how we can describe conditional statements in terms of projective limits, let us give an example. Suppose that we know that for a measuring device I, we have the following valid conditional statement that connects three outcomes $u, v, w \in X_I$: if for some object, the outcomes u and v are both possible, then the outcome w is also possible for this object.

We already know that if the outcome $u \in X_I$ is possible, this means that the set X of all possible values of the measured quantity (= projective limit) should contain a value a for which $a_I = v$. We also know that if the outcomes $u, v \in X_I$ are compatible ($u \sim v$), this means that there exists indistinguishable values $a, b \in X$ for which $a \sim b$, $a_I = u$, and $b_I = v$.

It is desirable to describe conditional statements like the above statement in the same terms. Let us show, for example, how the above statement should look like.

In our terms, an object with the value $a \in X$ of the measured quantity, the measuring device I *actually produces* an outcome $u \in X_I$ if $a_I = x$. It is natural to say that an object with the value $b \in X$ of the measured quantity *can produce* an outcome $u \in X_I$ if there exists an object $a \in X$ that is indistinguishable from b and for which $a_I = u$.

Correspondingly, an object with the value $c \in X$ of the measured quantity *can produce* outcomes $u \in X_I$ and $v \in X_I$ if there exist objects $a, b \in X$ that are indistinguishable from each other and from c and for which $a_I = u$ and $b_I = v$. The above conditional statement means that for any such object, w is also possible, i.e., there exists an object d that is indistinguishable from a, b, and c, for which $d_I = w$.

In this manner, we can describe an arbitrary conditional or disjunctive conditional statement in terms of the projective limit.

Comment. The above definition is rather complicated – it requires that we check a certain property for all possible pairs of values (infinite functions). Because of this complexity, this property is difficult to check.

It is therefore desirable to look for a simpler reformulation of the above idea, a reformulation that would enable us to check the validity of conditional statements (more) efficiently.

Examples: a brief introduction. Now we know how, in the general case, we can describe the set of all possible values of the measured quantity in terms of the corresponding family of measuring devices. Let us illustrate this general knowledge on the examples for which we already know how to describe the corresponding measuring devices.

Example 1: interval uncertainty leads to real numbers. Interval computations correspond to measuring real numbers. Let us show that by considering the corresponding family of measuring instruments, we can indeed get real numbers as the set of all possible values of the corresponding physical quantity.

In other words, what we plan to show is that there is no need to specifically require real numbers – real numbers naturally appear as a result of our analysis of a measurement process.

Indeed, each interval-related corresponding measuring instrument I can be described as a covering of the real line by several "infinite" intervals $(-\infty, \overline{x}]$ and $[\underline{x}, \infty)$, and several finite (bounded) intervals $[\underline{x}, \overline{x}]$. In real-life measurements, the endpoints a and b are usually rational numbers. The outcomes $v_1, \ldots, v_m \in X$ are compatible if and only if the corresponding intervals $S_1, \ldots, S_m$ intersect.

We can form a measuring device I by combining several measuring instruments $I', \ldots, I''$ of this type. In this case, possible outcomes of I are sequences $v = (v'_i, \ldots, v''_j)$ (where $v'_i \in X', \ldots, v''_j \in X''$), for which the corresponding intervals $S'_i, \ldots, S''_j$ intersect, and the corresponding sets S are the intersections of these intervals.

Since a non-empty intersection of two intervals is also an interval, such measuring devices have exactly the same form as the original measuring instruments: finitely many outcomes, and to each outcome, we put into correspondence an interval (finite or infinite) in such a way that these intervals cover the entire real line.

In interval terms, a *projection* relation $\pi(v) = v'$ means that the interval corresponding to S is contained in the interval S' corresponding to v': $S \subseteq S'$.

If we want to be able to describe the corresponding measurements with an arbitrary accuracy throughout the entire scale, we must require that for arbitrary bounds and for an arbitrary accuracy, there exists a measuring device in our family that measures all the values within the given bounds with a given accuracy.

Formally, it is sufficient to require that for every natural number n, there exist a measuring device I_n in which all the values from $-n$ to n are covered by intervals of length $\le 1/n$ and by none of the infinite intervals. Since we want a family of measuring instruments – in the above sense – we can, without losing generality, require that $I_n \sqsubseteq I_{n+1}$, i.e., that there is a projection $\pi_{n+1,n}$ from I_{n+1} to I_n.

What is the projective limit of this family of measuring devices? By definition, a value of the measured quantity is represented by a sequence of outcomes $a_n \in X_n$ such that each outcome a_n is the projection of the following outcome a_{n+1}. As we have just mentioned, the fact that a_n is a projection of a_{n+1} means that the interval S_{n+1} corresponding to a_{n+1} is a subset of the interval S_n corresponding to a_n.

Let us first consider the case when for some m, we get a finite interval S_m. Since the interval S_m is finite, its width is $\le 1/m$. The subsequent intervals $S_{m+1}, \ldots, S_n, \ldots$ are subsets of S_m, so they cannot be infinite. Thus, we get a sequence of nested intervals $S_m \subseteq S_{m+1} \subseteq \ldots \subseteq S_n \subseteq \ldots$ of width $w(S_n) \le 1/n \to 0$ as $n \to \infty$. Such nested sequence of intervals always contains exactly one real number. We will denote the real number corresponding to the sequence a by $v(a)$. For each n, we then have $v(a) \in S_n$.

If two sequences a and a' describe the same real number $v(a) = v(a')$, then for every n, the corresponding intervals S_n and S'_n have a common point – the point $v(a) = v(a')$. Thus, these intervals intersect, i.e., $a_n \sim_n a'_n$ for all n. According to our definition, this means that the values a and a' are indistinguishable, i.e., that $a \sim a'$.

Vice versa, if two sequences a and a' are indistinguishable ($a \sim a'$), this means that for each n, the corresponding intervals S_n and S'_n intersect. Since $v(a) \in S_n$, $v(a') \in S'_n$, and the widths of two intersecting intervals S_n and S'_n are $\le 1/n$, we thus conclude that $|v(a) - v(a')| \le 2/n$. Since this inequality must be true for all n, we have $v(a) = v(a')$.

Thus, on the projective limit, the compatibility relation is transitive – hence, an equivalence relation, and the equivalence classes corresponding to the finite-interval case are in 1-1 correspondence with the real numbers.

To complete our description, we must consider the remaining case, when all the intervals S_n are infinite. In this case, also $S_m \supseteq S_n$ for $m < n$. It is easy to see that an interval $(-\infty, \overline{x}]$ cannot contain an interval $[\underline{x}, \infty)$ and vice versa. So, in this case, we have only two possibilities: either all the intervals S_n are of the type $(-\infty, \overline{x}_n]$, or they are all of the type $[\underline{x}_n, \infty)$.

For the first possibility, since we assumed that for I_n, no infinite interval covers $[-n, n]$, we conclude that $\overline{x}_n < -n$. Thus, in this case, the actual value

of the measured quantity is smaller than $-n$ for every n. Intuitively, this value corresponds to $-\infty$.

This possibility is real: we can choose, for every n, the outcome a_n that covers $-\infty$.

Similarly, for the second possibility, since we assumed that for I_n, no infinite interval covers $[-n, n]$, we conclude that $\underline{x}_n > n$. Thus, in this case, the actual value of the measured quantity is larger than n for every n. Intuitively, this value corresponds to ∞. This possibility is also real: we can choose, for every n, the outcome a_n that covers $+\infty$.

There may be several sequences a, a' corresponding to $-\infty$. In each of these sequences, we have $S_n = (-\infty, \overline{x}_n]$ and $S'_n = (-\infty, \overline{x}'_n]$, so for every n, we have $S_n \cap S'_n \neq \emptyset$ hence $a_n \sim_n a'_n$. Thus, every two such sequences are indistinguishable.

Similarly, there may be several sequences a and a' corresponding to ∞; every two such sequences are indistinguishable.

It is easy to see that a finite set of sequences is indistinguishable if and only every two sequences are indistinguishable. Thus, we arrive at the following conclusion.

Conclusion. For the corresponding projective limit, indistinguishability is an equivalence relation. Equivalence classes with respect to this relation represent:

- all real numbers, and
- the two additional values $-\infty$ and $+\infty$.

Comment about relations and operations with real numbers. In measurements, when we represent measurement results by real numbers, we often actively use standard relations and operations that are defined on real numbers, such as the ordering relation $<$, arithmetic operations such as addition $+$, etc.

As of now, we have only shown that we get real numbers in the limit. The description of how the standard relations and operations look like in these new terms will be given after we explain, in general, how to describe relations and functions on projective limits.

Comment about infinite values. We have shown that, for interval uncertainty, the projective limit consists of all real numbers plus the additional values $-\infty$ and $+\infty$.

These additional infinite values make some *physical* sense: for example, some scientists believe that the physical Universe is infinite, and thus, e.g., its total energy is infinite – meaning that no matter what measuring device we use, we will always get an "above the scale" outcome.

These additional values also make perfect *computational* sense: G. W. Walster and others have shown (see, e.g., [42, 83]) that if we extend real numbers by infinite values, we get better interval estimations. Indeed, one way to find an enclosure for the range $\mathbf{y}$ of an algorithmically given function $f(x_1, \ldots, x_n)$

on given intervals $\mathbf{x}_1, \ldots, \mathbf{x}_n$ is to perform the so-called straightforward interval computations. In this technique, we simply replace, in the sequence of elementary operations that constitute the algorithm f, each operations with real numbers by the corresponding operation of interval arithmetic.

It is known that in general, we thus get an enclosure $\mathbf{Y} \supseteq \mathbf{y}$ for the desired range $\mathbf{y}$. It is also known that when the expression for $f(x_1, \ldots, x_n)$ is a *single-use expression* (SUE), i.e., if each variable occurs only once, then straightforward interval computations lead to the exact estimate $\mathbf{Y} = \mathbf{y}$.

In some case, an algebraic transformation can make the function that was previously not given in a SUE form into a SUE form – and thus, help interval computations. For example, $x_1 \cdot x_2 + x_1 \cdot x_3$ is not SUE, but it can transformed into an equivalent SUE form $x_1 \cdot (x_2 + x_3)$.

As another example of such a situation, we can consider a function $f(x_1, x_2) = x_1/(x_1 + x_2)$. This function is frequently used in real-life computations: e.g., if we know the probability x_1 of the outcome v_1 and the probability x_2 of the outcome v_2, then the conditional probability of v_1 under the condition that we have observed either v_1 or v_2 is equal exactly to $x_1/(x_1 + x_2)$. In this expression, the variable x_1 occurs twice so this function is not SUE. It can be transformed into a SUE form if we divide both numerator and denominator by x_1. As a result, we get the following SUE equivalent form: $1/(1 + x_2/x_1)$. However, if the interval of possible values of x_1 contains 0 – e.g., if $\mathbf{x}_1 = [0, 1]$ and $\mathbf{x}_2 = [1, 2]$ – the value of x_2/x_1 can be infinite, and infinite values are not covered by standard interval arithmetic. Thus, to use this SUE reformulation, we must extend the set of real numbers (and corresponding intervals) to infinite values as well.

One may ask: this example and the corresponding extension have already been proposed by Walster et al., so what do we gain by repeating it in our terms? Good news is that Walster et al. came up with the infinite values in an *ad hoc* way, as a way to improve estimations of interval uncertainty for real-life expressions. What we have just shown is that infinite values *naturally appear* in the description of real numbers if go deep enough in the analysis of how real numbers come from measurements.

Indeed, the inclusion of infinite values follows the same logic as the above definition of all possible sequences in a projective limit. Specifically, if there are no restrictions on the value of the physical quantity, then it is quite possible that for every measuring device, we get an "above scale" outcome. Thus, on every state of measurement, there is no possibility to exclude values that are actually infinite – and since there is no (and cannot be any) empirical reason to exclude such values, we do include them.

From this viewpoint, adding infinite values is as natural as adding irrational values – the example that we used before to explain this idea. So, if we do not object to the inclusion of all irrational values (and most physicists do not object to that), we should also include infinite values.

Example 2: counting leads to natural numbers. As we have mentioned, all real-life counting devices have a upper bound n so that:

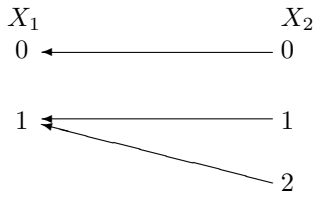
- all the values from 0 to $n-1$ are counted exactly, while
- any larger number of objects (i.e., n or larger) result in the same outcome.

As we have mentioned, for prehistoric people, this "$\geq n$" outcome meant "many". In the above text, we simply labelled it as n.

From this viewpoint, a real-life counting device does not always give us a full information about the number of objects. To get more information, we must use counting devices I_n with larger and larger bounds n.

How are projections defined in this case? If we know the result a_n of counting with a procedure I_n (that has a bound n), then, for every $n' < n$, we can uniquely reconstruct the result of $a_{n'}$ of applying the counting procedure $I_{n'}$ (that has a bound n') to the same object: namely, $a_{n'} = \min(a_n, n')$. So, in this case, $\pi_{n,n'}(a) = \min(a, n')$.

Here is a geometric illustration of this projection for $n = 2$ and $n' = 1$:



We have also mentioned that in this case, no two outcomes are compatible. What is the resulting projective limit?

- If for some n, the outcome $v \in X_n$ is below the corresponding bound ($v < n$), then it stays this way for all further n. In this case, the mapping a takes the form $a_n = \min(v, n)$. Clearly, this mapping represents the natural number v.
- On the other hand, if for every counting procedure I_n, we get the bound n, then this mapping $a_n = n$ also satisfies the projective property. This mapping represents (positive) infinity ∞.

Thus, for counting, the projective limit consists of all natural numbers plus the additional value ∞.

Does this value make physical sense? Absolutely. For quite some time, scientists believed that the physical Universe contained infinitely many particles (and some still believe this). In this physically possible case, no matter how many times you count, you always get the outcome "many" – this is exactly what infinity is about.

Example 3: "yes"-"no" measurements lead to truth values. In a single "yes"-"no" measurement, we may get 0 (= "no"), we may get 1 (= "yes"), and we may get U (unknown).

As we have mentioned, a typical example of "yes"-"no" measurement is the case when we measure a physical quantity x with an accuracy Δ and, based on the result $\widetilde{x}$ of this measurement, try to decide whether $x > x_0$. Then:

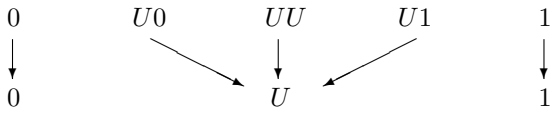
- When $\widetilde{x} \le x_0 - \Delta$, we can conclude that $x \le x_0$ and the answer is 0.
- When $\widetilde{x} > x_0 + \Delta$, we can conclude that $x > x_0$ and the answer is 1.
- When $x_0 - \Delta < \widetilde{x} \le x_0 + \Delta$, it is possible that $x \le x_0$ and it is also possible that $x > x_0$, so the answer is U.

Based on this example, it is clear that to get a more complete answer, we must repeat this procedure with a more accurate measuring device, for which the error bound Δ is smaller.

For the measuring device I_2 combining the two measuring instruments, possible outcomes are thus 0, 1, $U0$, $U1$, and U^2 (meaning U repeated twice). In the first four cases, we know the answer; in the case U^2, we must employ a yet more accurate measuring device, etc.

For a measuring device I_n consisting of n measuring instruments, the set X_n of possible outcomes consists of 0, 1, $U0$, $U1$, U^20, U^21, ..., $U^{n-1}0$, $U^{n-1}1$, and U^n. When $n > n'$, a projection $\pi_{n,n'}$ simply means that for any sequence of length $> n'$, we only keep the first n' measurement results (= symbols in a sequence).

For example, for $n = 2$ and $n' = 1$, the projection retains 0 and 1, and maps $U0$, $U1$, and UU into U:



What is the resulting projective limit?

- If for some n_0, the outcome ends in 0 or 1, i.e., is of the form $U^{n_0-1}0$ or $U^{n_0-1}1$, then it stays this way for all further $n > n_0$. In this case, the mapping a takes the following form:
  - for $n < n_0$, $a_n = U^n$;
  - for $n \ge n_0$, we have $a_n = U^{n_0-1}0$ (or $a_n = U^{n_0-1}1$).

  We will denote these cases by, correspondingly, $U^{n_0-1}0$ or $U^{n_0-1}1$.
- On the other hand, if for every measuring device I_n, we get the uncertain outcome U^n, then the corresponding mapping $a_n = U^n$ also satisfies the projective property. We will denote this mapping as U^∞.

Thus, for "yes"-"no" measurements, the projective limit X consists of the sequences of the type U^n0 and U^n1, where n is an arbitrary natural number, plus the additional value U^∞.

In physical terms, this limit X describes possible actual truth value of the statement that we are testing:

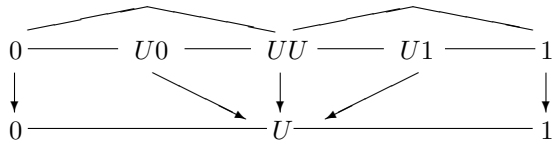
- sequences ending in 0 mean "false";
- sequences ending in 1 mean "true"; and
- the sequence U^∞ means "unknown".

What is the compatibility structure on this set X? In the previous section, we described the graph structure on the sets X_1 and X_2 corresponding to one and two sequential "yes"-"no" measurements. For arbitrary n, the graph structure on X_n is similar:

- all the values U^k0 ending in 0 are compatible with each other,
- all the values U^k1 ending in 1 are compatible with each other, and
- the value U^n is compatible with every other value.

It is easy to see that for X_n, the simplicial complex structure can be reduced to a graph structure: a set $S \subseteq X_n$ of physically possible outcomes is compatible if and only if every pair of outcomes from S is compatible.

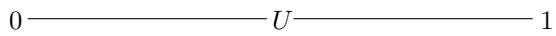
The above-defined projection $\pi_{n,n'}$ clearly preserves this compatibility relation. For example, for $n = 2$ and $n' = 1$, this projection takes the following form:



It is easy to see that the above definition leads to the following indistinguishability structure $\sim$ on the set X:

- all the values ending in 0 are indistinguishable from each other,
- all the values ending in 1 are indistinguishable from each other, and
- the value U^∞ is indistinguishable from every other value.

In other words, all the values $v \in X$ ending in 0 are equivalent to each other, so they can be viewed as a single truth value 0. Similarly, all the values $v \in X$ ending in 1 are equivalent to each other, so they can be viewed as a single truth value 1. The value U^∞ is indistinguishable with 0 and indistinguishable with 1, and 0 and 1 are *not* indistinguishable from each other. In short, if we combine all the values corresponding to "false" and all the values corresponding to "true", the graph $\langle X, \sim \rangle$ describing possible truth values looks exactly like the graph $\langle X_1, \sim_1 \rangle$ that describes a single "yes"-"no" measurement:



It is easy to see that for the limit set X, the simplicial complex structure can be reduced to a graph structure: a finite set S is indistinguishable if and only if every pair of outcomes from S is indistinguishable.

Example 4: a combination of several independent physical quantities. From the measurement viewpoint, what does it mean to have m independent physical quantities? A natural interpretation is that, for an arbitrary degree of accuracy, we can have m independent physical instruments for measuring these quantities.

We already know what it means for a measuring device I to constitute a combination of m independent measuring instruments $I_1, \ldots, I_m$: its set X of physically possible outcomes is a Cartesian product $X_1 \times \ldots \times X_m$ of the sets corresponding to individual measuring instruments, its compatibility relation $\sim$ has the form $(v_1, \ldots, v_m) \sim (v'_1, \ldots, v'_m) \leftrightarrow v_1 \sim_1 v'_1 \,\&\, \ldots \,\&\, v_m \sim_m v'_m$, etc.

To complete the description of the independent case, we must also explain what it means for a projection π from one device $X = X_1 \times \ldots \times X_m$ to another device $X' = X'_1 \times \ldots \times X'_m$ to reflect independence. A natural definition is that we should have independent projections for each of m instruments, i.e., that the function π should have the form $\pi(v_1, \ldots, v_n) = (\pi_1(v_1), \ldots, \pi_m(v_m))$.

If for some family of measuring devices, all devices and all projections correspond to m independent quantities, then, for each i from 1 to m, the measuring instruments corresponding to i-th component also form a family of measuring devices, with projective limits $\langle X_i^l, \sim_i^l \rangle$.

In this case, the projective limit $\langle X^l, \sim^l \rangle$ of the combined measuring instruments is a Cartesian product of the projective limits corresponding to m quantities, i.e., $X^l = X_1^l \times \ldots \times X_m^l$ and

$$(v_1, \ldots, v_m) \sim^l (v'_1, \ldots, v'_m) \leftrightarrow v_1 \sim_1^l v'_1 \,\&\, \ldots \,\&\, v_m \sim_m^l v'_m$$

(and similarly, for indistinguishable sets).

Comment. This result can be reformulated as follows. We have two approaches to handling the situation of m independent quantities:

- First, we can consider measuring instruments measuring each of m quantities. For each of these quantities, we form a projective limit X_i^l, and then combine these projective limits into a single set of values $X_1^l \times \ldots \times X_m^l$.
- Alternatively, on each step, we can combine the measuring instruments into a single measuring device, and form a projective limit of such measuring devices.

What we are saying is that these two approaches lead to exactly the same description $\langle X^l, \sim^l \rangle$.

We can describe this fact graphically by saying is that the following diagram is commutative:

$$\begin{array}{ccc} I_1, \ldots, I_m & \longrightarrow & I = I_1 \times \ldots \times I_m \\ \downarrow \quad \downarrow & & \downarrow \\ X_1^l, \ldots, X_m^l & \longrightarrow & X^l = X_1^l \times \ldots \times X_m^l \end{array}$$

4 Properties of Physical Quantities

4.1 A useful auxiliary result: we can always restrict ourselves to a sequence of measuring devices.

From the physical viewpoint, it is important to consider the most general families of measuring devices. When we defined a physical quantity, we mentioned that in some cases, we only have a simple sequence of measuring devices, the next one more informative than the previous one $I \sqsubseteq I'$. However, in many other real-life situations, we may have several different measuring devices corresponding approximately to the same level of accuracy, devices none of which is more informative than the other one.

Since the main objective of this paper is to provide the most general description of validated uncertainty, in our definition of a quantity, we considered the most general case of a family that is not necessarily linearly ordered by the relation $\sqsubseteq$.

From the purely mathematical viewpoint (of proving results), it is desirable to consider the simpler case of sequences. We will show that from the purely mathematical viewpoint, we can always restrict ourselves to the case when devices are linearly ordered. This restriction will simplify some of the proofs.

Let us first give the corresponding definition.

Definition. By a *sequence of measuring devices*, we mean a pair $\langle \{I_i\}_{i=0,1,\ldots}, \{\pi_{i+1,i}\}\rangle$, where I_i are measuring devices and each function $\pi_{i+1,i} : I_{i+1} \to I_i$ is a projection of measuring devices.

Comment. Based on these projections, we can define, for every $i < j$, a projection $\pi_{j,i} : I_j \to I_i$ as a composition

$$\pi_{j,i} = \pi_{j,j-1} \circ \ldots \circ \pi_{i+1,i}.$$

Proposition. From every family of measuring devices $\langle \mathcal{I}, \mathcal{P}\rangle$, we can select a sequence for which the following three properties hold:

- the set of values X corresponding to the family is in 1-1 correspondence with the set of values corresponding to the sequence,
- for each $a, b \in X$, $a \sim b$ in the sense of the family if and only if $a \sim b$ in the sense of the sequence, and
- for each finite set $S \in X$, S is indistinguishable in the sense of the family if and only if it is indistinguishable in the sense of the sequence.

Comment. In other words, from the viewpoint of defining and analyzing the physical quantity, it is sufficient to consider only measuring devices from the selected sequence.

Proof. This proof will use some of the ideas used in the above proof that an arbitrary element $v \in X_I$ is possible for some $a \in X$.

1°. Since the family $\mathcal{I}$ is denumerable, we can enumerate all its elements into a sequence $\mathcal{I} = \{J_1, J_2, \ldots, J_n, \ldots\}$. Let us first show that we can always select a sequence $I_1 \stackrel{\text{def}}{=} J_{i_1}$, $I_2 \stackrel{\text{def}}{=} J_{i_2}$, etc., that has the following two properties:

- first, each device in this sequence is less informative that the next device, i.e., $I_i \sqsubseteq I_{i+1}$, and
- second, each device $I \in \mathcal{I}$ is less informative than one of the devices from this sequence: $I \sqsubseteq I_i$.

We will select elements of this sequence one by one. For $j = 1$, we select, e.g., $i_1 = 1$. Let us now assume that we have already selected the elements $i_1, \ldots, i_j$, and let us show how we can select an element i_{j+1}. Let k be the first integer for which I_k is not less informative than one of the selected devices $J_{i_1}, \ldots, J_{i_j}$. Since the family of measuring devices is a directed set, there exists a measuring device J_m from this family that is at least as informative as both J_{i_j} and I_k: $J_{i_j} \sqsubseteq J_m$ and $J_k \sqsubseteq J_m$. We then take $i_{j+1} \stackrel{\text{def}}{=} m$.

Due to our construction, we have $J_{i_j} \sqsubseteq J_{i_{j+1}}$ for all j. We also have $J_k \sqsubseteq J_{i_{j+1}}$. Let us prove that for every measuring device J_l, there exists a j for which $J_l \sqsubseteq J_{i_j}$. Indeed, at each stage of this algorithm, we took k as the smallest value for which J_k is not yet less informative than one of the selected devices J_{i_j}, and selected a new device that makes k less informative than one of the selected devices. Thus, at each stage, we increase the smallest value k for which J_k is not yet covered by one of the selected devices, at least by 1. Thus, eventually, every device J_l will be indeed covered. The selection is completed.

2°. Due to the selection, each device in the selected sequence is indeed less informative than the next one, hence this sequence is a sequence of measuring devices in the sense of the above definition.

3°. Let us now show that the set of values of the physical quantity for this sequence is indeed in 1-1 correspondence with the set of values corresponding to the entire original family. In other words, we need to prove that the corresponding projective limits are in 1-1 correspondence.

Every element in a projective limit corresponding to the family is a mapping a that maps every device I from this family into an element $a_I \in X_I$ in such a way that a projection property $\pi_{I,J}(a_I) = a_J$ holds. If we restrict this map to devices from the selected sequence, we thus get a mapping corresponding to the sequence.

Let us show that, vice versa, once we have a mapping $a = \{a_1, a_2, \ldots, \}$ corresponding to the selected sequence (i.e., if we know the values $a_{I_i} = a_i$), then we can uniquely extend a to the entire family $\mathcal{I}$ of measuring devices. Indeed, according to the second property of the selected sequence, for every measuring device $I \in \mathcal{I}$, there exists an i for which $I \sqsubseteq I_i$, i.e., for which the family $\mathcal{P}$ contains a projection $\pi_{I_i,I} : I_i \to I$. We can then define a_I as $\pi_{I_i,I}(a_i)$. One can easily show that due to properties of projection, thus defined value a_I does not depend on which exactly i we choose, and the corresponding mapping $I \to a_i$ indeed satisfies the projection property.

Thus, values corresponding to the family are indeed in 1-1 correspondence with the values corresponding to the sequence.

4°. Let us now prove that the corresponding compatibility relation is the same.

Indeed, if $a \sim b$ in the sense of the family, this means that $a_I \sim_I b_I$ for all measuring devices I, in particular, for devices from the selected family. Thus, we can conclude that $a_i \sim_i b_i$ for all i, i.e., that $a \sim b$ in the sense of the sequence as well.

Vice versa, let us assume that a and b are indistinguishable in the sense of the sequence, i.e., $a_{I_i} \sim_{I_i} b_{I_i}$ for all i. Let us then prove that in this case, a and b are also indistinguishable in the sense of the family, i.e., $a_I \sim_I b_I$ for every I. Indeed, due to the second property of the selected sequence, for every I, there exists an i for which $I \sqsubseteq I_i$, i.e., for which $\mathcal{P}$ contains a projection $\pi_{I_i,I}$. We know that $a_{I_i} \sim_{I_i} b_{I_i}$. By applying the projection $\pi_{I_i,I}$ to a_{I_i} and b_{I_i} and by using the property that projection retains compatibility, we conclude that indeed $a_I \sim_I b_I$.

So, a pair is indistinguishable in the sense of family if and only if it is indistinguishable in the sense of the sequence.

5°. Similarly, we can proof that indistinguishability of sets of outcomes is the same whether we use families or sequences.

The proposition is proven.

4.2 When Is the Indistinguishability Relation $\sim$ Transitive: A Criterion

In some cases, the indistinguishability relation is transitive, in some cases, it is not. We have seen that both for measurements with interval uncertainty and for counting, the indistinguishability relation $\sim$ on the corresponding set of values X of the physical quantity is transitive. In such cases, the relation $\sim$ is an equivalence relation, so we can subdivide all the values into equivalence classes. For measurements with interval uncertainty, the equivalence classes are exactly real numbers plus two additional classes $-\infty$ and $+\infty$. For counting, the equivalence classes are exactly natural numbers plus an additional class ∞.

For the case of several independent quantities, we get a Cartesian product of the corresponding quantities $X = X_1 \times \ldots \times X_m$. One can easily check that when the relations $\sim_i$ corresponding to each X_i are transitive, then the relation corresponding to the Cartesian product is transitive as well. In particular, for several independent real-valued quantities, the corresponding set of equivalence classes is an extended m-dimensional vector space $(R \cup \{-\infty, +\infty\})^m$ that consists of all m-dimensional vectors with finite or infinite coordinates.

On the other hand, for truth values corresponding to "yes"-"no" measurements, the compatibility relation is not transitive: $0 \sim U$, $U \sim 1$, but $0 \not\sim 1$.

How to check transitivity: formulation of the problem. When we have transitivity, then the situation is easier: $\sim$ is an equivalence relation, so possible values of the physical quantity can be identified with equivalence classes. It is therefore desirable to check when $\sim$ is a transitive relation.

Directly applying the definition of transitivity is difficult, because this would mean testing this condition for all possible infinite sequences a, b, and c. It is therefore desirable to provide a more efficient criterion for testing transitivity.

To solve this problem, we will formulate and prove a Proposition describing a more efficient criterion, and then explain how this criterion can be applied in practice.

Proposition. Let $\langle \mathcal{I}, \mathcal{P} \rangle$ be a family of measuring devices $\langle X_I, \sim_I \rangle$, and let $\langle X, \sim \rangle$ be the corresponding projective limit. Then, the following two conditions are equivalent to each other:

- the compatibility relation $\sim$ is *transitive*, i.e., for all $a, b, c \in X$, if $a \sim b$ and $b \sim c$, then $a \sim c$;
- for every $I \in \mathcal{I}$, there exists a $J \in \mathcal{I}$ for which, if $a_J \sim_J b_J$ and $b_J \sim_J c_J$, then $a_I \sim_I c_I$.

Proof. 1°. Because of the result proven in the previous subsection, it is sufficient to consider the case when we have a *sequence* of measuring devices.

For a sequence, the second condition takes the following form:

For every i, there exists a j for which, if $a_j \sim_j b_j$ and $b_j \sim_j c_j$, then $a_i \sim_i c_i$.

2°. Let us first prove that the second condition implies the first one.

Indeed, let $a \sim b$ and $b \sim c$. We want to prove that $a \sim c$. By definition, $a \sim c$ means that $a_i \sim_i c_i$ for all i. Let us pick any i and prove that $a_i \sim_i c_i$. Indeed, due to the second condition, there exists a j for which $a_j \sim_j b_j$ and $b_j \sim_j c_j$ imply that $a_i \sim_i c_i$. By definition of $\sim$, the condition $a \sim b$ means that $a_k \sim_k b_k$ for all k, in particular, for $k = j$. Thus, we have $a_j \sim_j b_j$, and similarly, we have $b_j \sim_j c_j$, hence we can indeed conclude that $a_i \sim_i b_i$.

3°. Let us now prove that the first condition implies the second one.

We will prove it by reduction to a contradiction. Indeed, if no such j exists, this means that for every j, there exist values $a^{(j)}$, $b^{(j)}$, and $c^{(j)}$ for which $a_j^{(j)} \sim_j b_j^{(j)}$, $b_j^{(j)} \sim_j c_j^{(j)}$, but $a_i^{(j)} \not\sim_i c_i^{(j)}$. Based on this sequence of triples $(a^{(j)}, b^{(j)}, c^{(j)})$, we will construct mappings a, b, and c, for which $a \sim b$, $b \sim c$, but $a \not\sim c$ – specifically, $a_i \not\sim_i c_i$.

The construction consists of Steps 1, 2, ... On each Step k, we will have a set of triples $(a^{(j)}, b^{(j)}, c^{(j)})$ corresponding to all the values j from some infinite set of indices $J_k \subseteq N$. We start with the class of all such triple $J_0 = N$. At Step k, we define the values a_k, b_k, and c_k for this k, and further limit the class of the triples to the class J_k of all the triples from J_{k-1} that are consistent with the selected values a_k, b_k, and c_k in the sense that $a_k^{(j)} = a_k$, $b_k^{(j)} = b_k$, and $c_k^{(j)} = c_k$.

Let us now describe how exactly we can perform Step k. For each of the infinitely many triples $(a^{(j)}, b^{(j)}, c^{(j)})$, $j \in J_{k-1}$, the k-th projection $(a_k^{(j)}, b_k^{(j)}, c_k^{(j)})$ is an element of the finite set $X_k \times X_k \times X_k = X_k^3$ of all possible triples of outcomes of the device I_k. Thus, at least one of these triples is repeated infinitely many times in the sequence $(a^{(j)}, b^{(j)}, c^{(j)})$. We will pick one of such infinitely repeated triples as the triple (a_k, b_k, c_k). This way, we guarantee that the selected set J_k is still infinite.

Let us show that the resulting mappings a, b, and c indeed belong to the set X, i.e., indeed satisfy the projection property. It is sufficient to prove it for a. Indeed, on each Step k, we restrict ourselves to indices $j \in J_k$ for which $a_k^{(j)} = a_k$. Thus, on the next Step $k+1$, we only consider such values $a^{(j)}$. The value a_{k+1} selected on the next step is equal to $a_{k+1}^{(j)}$ for one of the indices $j \in J_k$. By definition of X, we have $\pi_{k+1,k}(a_{k+1}^{(j)}) = a_k^{(j)}$, hence $\pi_{k+1,k}(a_{k+1}) = a_k$.

4°. Let us show that for the corresponding elements, $a \sim b$ and $b \sim c$.

It is sufficient to prove it for a and b. We need to prove that for every k, we have $a_k \sim_k b_k$. Indeed, for every k, we selected the triple (a_k, b_k, c_k) as a triple that appear in a sequence $(a_k^{(j)}, b_k^{(j)}, c_k^{(j)})$ for infinitely many different indices $j \in J_{k-1}$. Since there are infinitely many such indices, at least one of these indices j must be $\geq k$. For this j, we have $a_k = a_k^{(j)}$ and $b_k = b_k^{(j)}$. By definition of a triple $(a^{(j)}, b^{(j)}, c^{(j)})$, we have $a_j^{(j)} \sim_j b_j^{(j)}$. Since $k \leq j$, we can thus conclude that $\pi_{j,k}(a_j^{(j)}) \sim_k \pi_{j,k}(b_j^{(j)})$.

Since $a^{(j)}$ and $b^{(j)}$ are elements of the projective limit, the value $\pi_{j,k}(a_j^{(j)})$ is equal to $a_k^{(j)}$, i.e., to a_k. Similarly, $\pi_{j,k}(a_j^{(j)}) = b_k$, so we indeed conclude that $a_k \sim_k b_k$ for all k.

5°. To complete the proof, we must now prove that $a_i \not\sim_i c_i$.

Indeed, according to our construction, there exists a j for which $a_i = a_i^{(j)}$, $b_i = b_i^{(j)}$, and $c_i = c_i^{(j)}$. By definition of the triples $(a^{(j)}, b^{(j)}, c^{(j)})$, for each of these triples, we have $a_i^{(j)} \not\sim_i b_i^{(j)}$, hence $a_i \not\sim_i c_i$.

The proposition is proven.

If indistinguishability is transitive, then from each i, we can effectively find j for which $a_j \sim_j b_j$ and $b_j \sim_j c_j$ imply $a_i \sim_i c_i$. The above proof is not constructive: it proves the existence of the desired j not by explicitly constructing such a j but by reduction to a contradiction (i.e., by showing that if such a j does not exist, then we get a contradiction).

We can show, however, that there is an *effective* way of finding such a j. First, it is easy to show that if this implication is true for $j < i$, it is thus true for all larger j as well. Thus, it is sufficient to consider the case when $j \geq i$.

In this case, the condition $a_i \sim_i c_i$ can be described in terms of a_j and c_j, as $\pi_{j,i}(a_j) \sim_i \pi_{j,i}(c_j)$.

For each j, there are only finitely many possible pairs of triples $(a_j, b_j, c_j) \in X_j^3$, and the relations $\sim_j$ and $\sim_i$ are algorithmically decidable. Therefore, by testing all possible triples, we can effectively check whether for this j, $a_j \sim_j b_j$ and $b_j \sim_j c_j$ indeed imply $a_i \sim_i c_i$.

If the implication is valid, we got our j. If not, we increase j by 1 and test again, etc., until we find the desired value j.

4.3 How to Extend this Result to General Results and Techniques: Topology, Compactness, etc.

Observations. In the above proof, we started with a sequence of triples $(a^{(j)}, b^{(j)}, c^{(j)})$, and from this sequence, we "extracted" a subsequence that was, in some reasonable set, "converging" to the final object.

In other words, in our proof, we did two things:

- first, we used a natural analogue of the notion of convergence, a notion that is well defined in topology and in classical analysis, and
- we also used the property that, crudely speaking, from every sequence we can extract a convergent subsequence; in classical analysis, this property is one of the equivalent formulations of compactness.

Since we may have to proof other results like that (one of such statements-to-prove is presented in the next subsection), it is desirable to explicitly formalize this similarity, and to explicitly define the notions such as convergence and compactness for projective limits.

One more reason why compactness seems to be a natural idea is that when we described, e.g., natural numbers, we get, in addition to all normal natural numbers, a special value ∞ – a "limit" of other values. For the set of all real numbers, we get two extra values. These values are usually added if we want to *compactify* the corresponding topological space. So maybe infinities will also look natural if we describe such compactness explicitly.

Notion of convergence. Intuitively, when we say that a is a limit of the sequence $a^{(n)}$, we mean that the larger n, the closer $a^{(n)}$ to a. In other words, this means whatever measuring instrument I we take, no matter how accurate I is, for sufficiently large n, we will not be able to distinguish $a^{(n)}$ and a.

Thus, we arrive at the following natural definition.

Definition. Let $\langle \mathcal{I}, \mathcal{P} \rangle$ be a family of measuring devices $I = \langle X_I, \sim_I \rangle$, and let $\langle X, \sim \rangle$ be the range of the corresponding measured quantity. We say that a sequence $a^{(n)}$ of elements from X *converges* to an element $a \in X$ if for every I, there exists an integer N_0 such for every $n \geq N_0$, we have $a_I^{(n)} \sim_I a_I$.

Proposition. For the above model of real numbers, the new definition of convergence coincides with the standard mathematical notion of convergence.

Proof. 1°. Let us first prove that the standard convergence implies the new convergence.

We will first consider the case when the limit a corresponds to the actual real number $v(a)$.

Due to our description of the measuring process corresponding to real numbers, k-th measuring device I_k measures all the values within $[-k, k]$ with a measurement error $\leq 1/k$ (in particular, all outcomes corresponding to this range correspond to finite intervals). Thus, for every integer $k > |v(a)| + 1$, k-th measurement result a_k corresponds to a finite interval $S_k = [\underline{x}_k, \overline{x}_k]$ of width $\leq 1/k$.

For the same k-th measuring device, the value $a_k^{(n)}$ is described either by the same interval, or by one of the (finitely many) other intervals corresponding to the device I_k. The only way for $a_k^{(n)}$ to be not compatible with a_k is when the corresponding intervals do not intersect. In the case of non-intersection, the interval $S_k^{(n)}$ corresponding to the outcome $a_k^{(n)} \in X_k$ can be either to the left, or to the right of the interval S_k.

If the non-intersecting interval from X_k is to the left of S_k, this means that its upper endpoint is $< \underline{x}_k$. Let x^- denote the largest of such endpoints. Since there are finitely many of them, we have $x^- < \underline{x}_k$. Similarly, if the non-intersecting interval is to the right of S_k, then its lower endpoint is $> \overline{x}_k$. Let x^+ be the smallest of such endpoints; here, $x^+ > \overline{x}_k$.

Let $\delta > 0$ denote the smallest of the two positive numbers $\underline{x}_k - x^-$ and $x^+ - \overline{x}_k$.

Since $v(a^{(n)}) \to v(a)$, there exists an N_0 after which $|v(a^{(n)}) - v(a)| \leq \delta$. The value $v(a)$ in within the interval S_k, hence $v(a^{(n)})$ is within a δ-vicinity of this interval. Due to our choice of δ, this vicinity cannot contain any points from any intervals from X_k that do not intersect with S_k – thus, the interval corresponding to $a_k^{(n)}$ must have a common intersection with S_k. Hence, we must have $a_k^{(n)} \sim_k a_k$.

For the case when a corresponds to an infinite value, the proof is similar.

2°. Let us now prove that the new convergence implies standard convergence.

Again, we will restrict our proof to the case when a corresponds to the real number; the infinite case is proven similarly.

For the real number case, when k is large enough, the intervals S_k and $S_k^{(n)}$ corresponding to the values a_k and $a_k^{(n)}$ are of width $\le 1/k$. So, when these intervals intersect, the values $v(a) \in S_k$ and $v(a^{(n)}) \in S_k^{(n)}$ can differ by more than $2/k$.

Thus, from the convergence in the new sense, we conclude that for every k, there exists an N_0 such that for every $n > N_0$, we have $|v(a^{(n)}) - v(a)| \le 2/k$ – in other words, that $v(a^{(n)}) \to v(a)$ in the normal sense.

The proposition is proven.

Compactness result. Let us prove that for this notion of convergence, the set X is indeed "compact" in the sense that from every sequence, we can extract a convergent subsequence:

Proposition. Let $\langle \mathcal{I}, \mathcal{P} \rangle$ be a family of measuring devices $I = \langle X_I, \sim_I \rangle$, let $\langle X, \sim \rangle$ be the range of the corresponding measured quantity, and let $a^{(n)}$ be a sequence of elements from X. Then, there exists a subsequence $a^{(j_i)}$ and an element $a \in X$ such that $a^{(j_i)}$ converges to a (in the sense of the above definition).

Proof. In view of the first result from this section, it is sufficient to consider the case when we have a sequence of measuring devices.

The construction consists of Steps 1, 2, ... On each Step k, we will have a set of values $a^{(j)}$ corresponding to all the values j from some infinite set of indices $J_k \subseteq N$. We start with the class of all such values $J_0 = N$. At Step k, we define the value a_k for this k, the value $j_k > j_{k-1}$, and further limit the class J_k of all the elements from J_{k-1} that are consistent with the selected value a_k in the sense that $a_k^{(j)} = a_k$.

Let us now describe how exactly we can perform Step k. For each of the infinitely many values $a^{(j)}$, $j \in J_{k-1}$, the k-th projection $a_k^{(j)}$ is an element of the finite set X_k of all possible outcomes of the device I_k. Thus, at least one of these outcomes is repeated infinitely many times in the sequence $a^{(j)}$. We will pick one of such infinitely repeated outcomes as the outcome a_k. (This way, we guarantee that the selected set J_k is still infinite.) We want to select one of corresponding values j as j_k. Since there are infinitely many such js, we can select the first $j > j_{k-1}$ with this property.

Let us show that the resulting mapping a indeed belongs to the set X, i.e., that it indeed satisfies the projection property. Indeed, on each Step k, we restrict ourselves to triples $j \in J_k$ for which $a_k^{(j)} = a_k$. Thus, on the next Step $k+1$, we only consider such values $a^{(j)}$. The value a_{k+1} selected on the next

step is equal to $a_{k+1}^{(j)}$ for one of the indices $j \in J_k$. By definition of X, we have $\pi_{k+1,k}(a_{k+1}^{(j+1)}) = a_k^{(j+1)}$, hence $\pi_{k+1,k}(a_{k+1}) = a_k$.

Let us show that $a^{(j_k)} \to a$. Indeed, in this case, for every k, we can choose $N_0 = k$; for all further n, due to our construction, we have $a_k^{(j_n)} = a_k$, hence $a_k^{(j_n)} \sim_k a_k$.

Discussion: possible applications to the inverse problem. We have proven, in effect, that from the viewpoint of a natural measurement-relation notion of convergence, every set of values of a measured quantity is, in some reasonable sense, compact.

Of course, there exist non-compact sets, e.g., the set of all real numbers is not compact, the set of all integers is not compact, the set of all continuous functions is not compact. What we are claiming is that these sets are not compact because their traditional mathematical topologies, while mathematically natural and convenient, *do not* fully reflect the measurement origin of the corresponding physical quantities. If we reflect this origin fully, we get a compact set.

OK, we do not always have exactly a compact set, because – as for the set of truth values – the indistinguishability relation is not necessarily transitive. However, we have some important properties of a compact set that make our analysis easier.

Why is it important to have a compact set? One reason is that when we reconstruct the physical world from measurements, we solve the so-called *inverse problem*. A physical theory usually enables us to predict, given the actual state s of the world, the values m of the observable quantities. In other words, a physical theory usually provides us with a function T that maps each state of the world s into the corresponding value $m = T(s)$ of the observable quantities. This prediction is called a *forward problem*: given the state of the world, predict what will be observed.

Usually in physics, close states lead to similar values of observable quantities and the closer the two states, the close the corresponding observable values. In mathematical terms, it mean that the function T is continuous.

In practice, we observe the values of $m = T(s)$, and based on these values, we must reconstruct the state of the world s, i.e., we must solve the *inverse* problem: given $T(s)$ $(= m)$, reconstruct s.

When we perform enough measurements to reconstruct s more or less uniquely, this problem becomes the problem of applying the inverse function T^{-1} to the observed values m. The problem is that the inverse function is not necessarily continuous. As a result, small measurement errors in measuring m can lead to drastic errors in the reconstructed state $s = T^{-1}(m)$. Such problems are usually called *ill-posed*; see, e.g., [77].

There are numerous real-life examples of ill-posed problems ranging from image processing to signal reconstruction to determining the internal structure of the Earth from the results of seismic observations and geophysical experiments.

From this viewpoint, compactness is a perfect property because it guarantees that an inverse of every 1-1 continuous function is also continuous. Thus, hope-

fully, our general approach to validated uncertainty will be helpful in solving ill-posed inverse problems.

Additional topology-like definition: closed set. We can say that a set $S \subseteq X$ is *closed* if it contains all its limit points, i.e., $\{a^{(n)}\} \subseteq S$ and $a^{(n)} \to a$ implies that $a \in A$.

Comment. For real numbers, since our notion of convergence coincides with the standard one, this definition also coincides with the standard definition of a closed set.

Additional topology-like definition: open set. We can say that a set $S \subseteq X$ is *open* if for every element $a \in S$, there exists a measuring device I such that if $b_I \sim_I a_I$, then $b \in S$.

Comment. One can easily check that for real numbers, this definition also coincides with the standard definition of an open set.

Comment. It would be interesting to analyze these seemingly natural definitions in the general case. For example, if we define closed and open sets in this manner, is it always true that a complement to an open set is closed and vice versa?

4.4 When Can We Reduce Simplicial Complex Description to a Graph Description?

Problem. In some cases – e.g., for real numbers – a finite set of values is indistinguishable if and only if every pair from this set is indistinguishable. In some other cases – as the ones described before for measuring devices – this may not be true. It is desirable to find a simple criterion for checking when a simplicial complex description can be reduced to the graph description.

Definition. Let $\langle \mathcal{I}, \mathcal{P} \rangle$ be a family of measuring devices $\langle X_I, \mathcal{S}_I \rangle$, and let $\langle X, \mathcal{S} \rangle$ be the corresponding projective limit. For every measuring device $I \in \mathcal{I}$, we say that a finite set $S = \{a^{(1)}, \ldots, a^{(n)}\} \subseteq X$ is *I-compatible* if its I-th projections are compatible, i.e., if $\{a_I^{(1)}, \ldots, a_I^{(n)}\} \in \mathcal{S}_I$.

We say that a finite set $S = \{a^{(1)}, \ldots, a^{(n)}\} \subseteq X$ is *pairwise I-compatible* if its I-th projections are pairwise compatible, i.e., if $a_I^{(i)} \sim_I a_I^{(j)}$ (or, equivalently, $\{a_I^{(i)}, a_I^{(j)}\} \in \mathcal{S}_I$) for all $i, j \le n$.

Proposition. Let $\langle \mathcal{I}, \mathcal{P} \rangle$ be a family of measuring devices $\langle X_I, \mathcal{S}_I \rangle$, and let $\langle X, \mathcal{S} \rangle$ be the corresponding projective limit. Then, the following two conditions are equivalent to each other:

- a finite set of values is indistinguishable if and only if every pair of values is indistinguishable;
- for every natural number n and for every measuring device $I \in \mathcal{I}$, there exists a $J \in \mathcal{I}$ such that if every two elements from the n-element set $S = \{a^{(1)}, \ldots, a^{(n)}\}$ are pairwise J-compatible, then the entire set is I-compatible.

Comments. The proof of this proposition is similar to the proof of the proposition about transitivity. Similar to that proof, we can also algorithmically find J from n and I.

5 Functions between Physical Quantities

This section will form Part 2 of our text. In Part 2, we will cover the following three topics:

- how to describe functional relations between the physical quantities;
- how to describe relations and operations that form the structure of a physical quantity – such as the ordering of real numbers, addition of real numbers, etc.;
- how to describe and analyze situations in which the desired physical "quantity" is itself a function; examples of such situations are given, e.g., in [2, 16, 45, 81, 82].

6 Future Work

More efficient ways to take conditional knowledge into consideration when describing the measured quantity. In the above text, we have shown how, based on the information about measuring devices that measure a given quantity with higher and higher accuracy, we can design a natural representation of the measured quantity. The resulting description is reasonably efficient for the simpler cases: namely, for cases when for each measuring device, the only information that we have about different outcomes is whether these outcomes, when observed in two measurements, guarantee that the corresponding objects are different.

In real life, in addition to such information, we may have *conditional* information about the outcomes, i.e., information of the type "if, for some object, we have observed x_1 and x_2, then the outcome x_3 is also possible for this same object. In the above text, we have shown how to describe such statements, but the problem is that our description is not very efficient. It is therefore desirable to design more efficient reformulations of our definitions.

Extension to probabilistic uncertainty. In this paper, our objective was to come up with a general description of validated uncertainty. Validated uncertainty corresponds to the description, for every possible actual value of the measured quantity and for each measuring device, of all possible measurement results. Often, in addition to the set of possible measurement results, we have an additional information about the relative frequency (probability) of different results; see, e.g., [62]. It is desirable to extend our general description to this probabilistic case. Some work on extending domains to the probabilistic case has already been done; see, e.g., [22]. It is desirable to extend this work to our most general case.

Extension to the case when do not have a full theory. The above analysis of validated uncertainty started with the description of measuring devices, description in which we assumed that we have a theory that fully describes both the measuring device and the quantity that this device measures. In complex situations related to advanced measurements, we may not have a full theory of the measured quantity. It is therefore desirable to extend our description to the case when no such full theory is available.

Formalizing the notion of possibility. Our objective is to provide a formalism that is useful for engineering and applied physics. In engineering and applied physics, it is often important to know what is possible and what is not possible. It is therefore desirable to add a formalized notion of possibility to our general formalism. This can be potentially done by using modal logic [12, 28, 53, 65] – a part of logic that studies notions related to possibility.

Acknowledgments

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

This work was performed during V.K.'s visit to Brazil. This visit was sponsored by CTINFO/CNPq.

References

[1] O. Aberth, *Precise Numerical Analysis Using C++*, Academic Press, New York, 1998.

[2] M. S. Aguiar, A. C. da Rocha Costa, and G. P. Dimuro, "ICTM: An Interval Tessellation-Based Model for Reliable Topographic Segmentation", *Numerical Algorithms*, 2004 (to appear).

[3] A. Aho, J. Hopcroft, and J. Ulman, *Data Structures and Algorithms*, Addisson-Wesley, Boston, Massachusetts, 1983.

[4] J. Armengol, J. Vehi, L. Trave-Massuyes, and M. A. Sainz, "Application of Modal Intervals to the Generation of Error-Bounded Envelopes", *Reliable Computing*, 2001, Vol. 7, No. 2, pp. 171–185.

[5] J. Barwise (ed.), *Handbook of Mathematical Logic*, North-Holland, Amsterdam, 1977.

[6] M. J. Beeson, *Foundations of Constructive Mathematics*, Springer-Verlag, New York, 1985.

[7] G. Belfonte and B. Bona, "An improved parameter identification algorithm for signal with unknown-but-bounded errors", *Proc. 7th IFAC Symposium on Identificatio and Parameter Estimation*, York, UK, 1985.

[8] B. Bouchon-Meunier and V. Kreinovich, "From Interval Computations to Modal Mathematics: Applications and Computational Complexity", *ACM SIGSAM Bulletin*, 1998, Vol. 32, No. 2, pp. 7–11.

[9] E. Bishop, *Foundations of Constructive Analysis*, McGraw-Hill, 1967.

[10] E. Bishop and D. S. Bridges, *Constructive Analysis*, Springer-Verlag, New York, 1985.

[11] D. S. Bridges, *Constructive Functional Analysis*, Pitman, London, 1979.

[12] B. Chellas, *Modal Logic: An Introduction*, Cambridge University Press, Cambridge, 1980. Cresswell, M. J., 1995,

[13] F. L. Chernousko, *Estimation of the Phase Space of Dynamical Systems*, Nauka Publ., Moscow, 1988 (in Russian).

[14] F. L. Chernousko, *State Estimation for Dynamical Systems*, CRC Press, Boca Raton, Florida, 1994.

[15] L. Danzer, B. Grünbaum, and V. Klee, *Helly's theorem and its relatives*, *Proc. 1963 Sympos. Pure Math. – Convexity*, Amer. Math. Soc., Providence, Rhode Island, 1963, pp. 101–180.

[16] G. J. Deboeck, K. Villaverde, and V. Kreinovich, "Interval Methods for Presenting Performance of Financial Trading Systems", *Reliable Computing*, 1995, Supplement (Extended Abstracts of APIC'95: International Workshop on Applications of Interval Computations, El Paso, Texas, Febr. 23–25, 1995), pp. 67–70.

[17] G. P. Dimuro and A. C. da Rocha Costa, "Towards a domain-theoretic modelling of a measuring process", *Frontiers in Artificial Intelligence and its Applications*, 2001, Vol. 71, pp. 68–75.

[18] G. P. Dimuro and A. C. da Rocha Costa, "Towards a domain-theoretic modelling of sensor data fusion", *Frontiers in Artificial Intelligence and its Applications*, 2003, Vol. 101, pp. 103–110.

[19] G. P. Dimuro, A. R. da Rocha Costa, and D. M. Claudio, "A bi-structural coherence space for a global representation of the system IR of real intervals", *Trends in Information Technology, Proceedings of the International Conference on Information Technology*, Bhubaneswar, India, 1999, pp. 173–178. of rational

[20] G. P. Dimuro, A. R. da Rocha Costa, and D. M. Claudio, "A coherence space of rational intervals for a construction of IR", *Reliable Computing*, 2000, Vol. 6, No. 2, pp. 139–178.

[21] J. Eckhoff, "Helly, Radon, and Carathéodory Type Theorems", In: P. M. Gruber and J. M. Wills (eds.), *Handbook of Convex Geometry*, North-Holland, Amsterdam, 1993, pp. 389–448.

[22] A. Edalat, "Domains for computation in mathematics, physics, and exact real arithmetic", *Bulletin of Symbolic Logic*, 1997, Vol. 3, No. 4, pp. 401–452.

[23] H.B. Enderton, *A Mathematical Introduction to Logic*, Academic Press, New York, 1972.

[24] R. P. Feynman, R. B. Leighton, and M. Sands, *Feynman Lectures on Physics*, Addison-Wesley, Boston, Massachusetts, 1989.

[25] A. F. Filippov, "Ellipsoidal estimates for a solution of a system of differential equations", *Interval Computations*, 1992, No. 2(4), pp. 6–17.

[26] A. Finkelstein, O. Kosheleva, and V. Kreinovich, "Astrogeometry, error estimation, and other applications of set-valued analysis", *ACM SIGNUM Newsletter*, 1996, Vol. 31, No. 4, pp. 3–25.

[27] P. C. Fishburn, *Interval Orders and Interval Graphs*, Wiley, New York, 1985.

[28] M. Fitting and R. Mendelsohn, *First Order Modal Logic*, Kluwer, Dordrecht, 1998.

[29] E. Fogel and Y. F. Huang, "One the value of information in system identification. Bounded noise case", *Automatica*, 1982, Vol. 18, pp. 229–238.

[30] E. Gardenes, H. Mielgo, and A. Trepat, "Modal intervals: reason and ground semantics", In: K. Nickel (ed.), *Interval Mathematics 1985*, Lecture Notes in Computer Science, Vol. 212, Springer-Verlag, Berlin, Heidelberg, 1986, pp. 27–35.

[31] E. Gardenes, M. A. Sainz, L. Jorba, R. Calm, R. Estela, H. Mielgo, and A. Trepat, "Modal intervals", *Reliable Computing*, 2001, Vol. 7, No. 2, pp. 77–111.

[32] J.-Y. Girard, "The system F of variable types, fifteen years later", *Theoretical Computer Science*, 1986, Vol. 45.

[33] J.-Y. Girard, "Linear logic", *Theoretical Computer Science*, 1987, Vol. 50.

[34] J.-Y. Girard, P. Taylor, and Y. Lafont, *Proofs and Types*, Cambridge University Press, Cambridge, 1989.

[35] M. C. Golumbic, *Algorithmic Graph Theory and Perfect Graphs*, Academic Press, New York, 1980.

[36] C. A. Gunter and D. S. Scott, "Semantic domains", In: J. van Leeuwen, *Handbook of Theoretical Computer Science*, Elsevier, 1990, pp. 635–674.

[37] P. Halmos, *Lectures on Boolean Algebras*, Van Nostrand, Princeton, 1963.

[38] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control and Robotics*, Springer-Verlag, London, 2001.

[39] T. Jech, *Set Theory*, Springer-Verlag, Berlin, New York, 1997.

[40] R. B. Kearfott, *Rigorous Global Search: Continuous Problems*, Kluwer, Dordrecht, 1996.

[41] R. B. Kearfott and V. Kreinovich, *Applications of Interval Computations*, Kluwer, Dordrecht, 1996.

[42] R. B. Kearfott and H. Muñoz, "Slope Intervals, Generalized Gradients, Semigradients, Slant Derivatives, and Csets", *Reliable Computing* (to appear).

[43] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Applications*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[44] V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl, *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Kluwer, Dordrecht, 1999.

[45] V. Kreinovich, H. T. Nguyen, G. P. Dimuro, A. C. da Rocha Costa, and B. Bedregal, "A New Differential Formalism for Interval-Valued Functions and Its Potential Use in Detecting 1-D Landscape Features", *Proceedings of the International Conference on Information Technology InTech'03*, Chiang Mai, Thailand, December 17–19, 2003, pp. 491–498.

[46] V. Kreinovich and K.Villaverde, *Towards modal interval analysis: how to compute maxima and minimia of a function from pproximate measurement results*, University of Texas at El Paso, Computer Science Department, Technical Report UTEP-CS-91-9, 1991.

[47] B. A. Kushner, *Lectures on Constructive Mathematical Analysis*, Amer. Math. Soc., Providence, Rhode Island, 1984.

[48] S. Lefschetz, *Introduction to Topology*, Princeton University Press, Princeton, 1949.

[49] L. R. Lewis and C. H. Papadimitrious, *Elements of the Theory of Computation*, Prentice-Hall, Englewood Cliffs, New Jersey, 1981.

[50] S. Li, Y. Ogura, and V. Kreinovich, *Limit Theorems and Applications of Set Valued and Fuzzy Valued Random Variables*, Kluwer Academic Publishers, Dordrecht, 2002.

[51] J. Lorkowski and V. Kreinovich, "If we measure a number, we get an interval. What if we measure a function or an operator?", *Reliable Computing*, 1996, Vol. 2, No. 3, pp. 287–297.

[52] T. A. McKee and F. R. McMorris, *Topics in Intersection Graph Theory*, SIAM, Philadelphia, 1999.

[53] G. E. Mints, *A Short Introduction to Modal Logic*, Center for the Study of Language and Information (CSLI), Stanford University, Stanford, California, 1992.

[54] D. Misane and V. Kreinovich, "A new characterization of the set of all intervals, based on the necessity to check consistency easily", *Reliable Computing*, 1995, Vol. 1, No. 3, pp. 285–298.

[55] D. Misane and V. Kreinovich, "The necessity to check consistency explains the use of parallelepipeds in describing uncertainty", *Geombinatorics*, 1996, Vol. 5, No. 3, pp. 109–120.

[56] J. D. Monk and R. Bonnet (eds), *Handbook of Boolean Algebras*, North-Holland, Amsterdam, 1989.

[57] R. E. Moore, *Methods and Applications of Interval Analysis*, SIAM Publ., Philadelphia, 1979.

[58] H. T. Nguyen and V. Kreinovich, "Nested Intervals and Sets: Concepts, Relations to Fuzzy Sets, and Applications", In: R. B. Kearfott and V. Kreinovich, *Applications of Interval Computations*, Kluwer, Dordrecht, 1996, pp. 245–290.

[59] H. T. Nguyen and E. Walker, *First Course in Fuzzy Logic*, CRC Press, Boca Raton, Florida, 1999.

[60] J. P. Norton, "Identification and application of bounded parameter models", *Proc. 7th IFAC Symsposium on Identification and Parameter Estimation*, York, UK, 1985.

[61] M. Pour-El and J. Richards, *Computability in Analysis and Physics*, Springer-Verlag, New York, 1989.

[62] S. Rabinovich, *Measurement Errors: Theory and Practice*, American Institute of Physics, New York, 1993.

[63] R. H. S. Reiser, G. P. Dimuro, and A. C. da Rocha Costa, "A programming language for the Interval Geometric Machine", *Electronic Notes in Theoretical Computer Science*, 2003, Vol. 84.

[64] R. H. S. Reiser, G. P. Dimuro, and A. C. da Rocha Costa, "The Interval Geometric Machine model", *Numerical Algorithms*, 2004 (to appear).

[65] S. Reyes and M. Clarke, *Logic for Computer Science*, Addison-Wesley, Boston, Massachusetts, 1990.

[66] J. R. Schoenfield, *Mathematical Logic*, Addison-Wesley, Boston, Massachusetts, 1967.

[67] F. C. Schweppe, "Recursive state estimation: unknown but bounded errors and system inputs", *IEEE Transcations on Automatic Control*, 1968, Vol. 13, p. 22.

[68] F. C. Schweppe, *Uncertain Dynamics Systems*, Prentice Hall, Englewood Cliffs, New Jersey, 1973.

[69] D. S. Scott, "Data types as lattices", *SIAM J. Comput.*, 1976, Vol. 5, No. 3, pp. 522–587.

[70] D. S. Scott, "Logic and programming languages", *Communications of the ACM*, 1977, Vol. 20, No. 9, pp. 634–641.

[71] D. S. Scott, "Lambda calculus: some models, some philosophy", In: J. Barwise, H. J. Keisler, and K. Kunen (eds.), *The Kleene Symposium*, North-Holland, Amsterdam, 1980, pp. 223–265.

[72] D. S. Scott, "Domains for denotational semantics", In: M. Nielsen and E. M. Schmidt (eds.), *Automata, Languages, and Programming, 9th Symposium*, Springer Lecture Notes in Computer Science, Springer-Verlag, Berlin, 1982, Vol. 140, pp.577–613.

[73] D. S. Scott, "Some ordered sets in computer science", In: I. Rival (ed.), *Ordered Sets*, 1982, pp.677–718.

[74] D. S. Scott, "Lectures on a mathematical theory of computation", In: *Theoretical Foundations of Programming Methodology*, D. Reidel Publ., 1982, pp. 145–292.

[75] S. T. Soltanov, "Asymptotic of the function of the outer estimation ellipsoid for a linear singularly perturbed controlled system", In: S. P. Shary and S. P. Shokin (eds.), *Interval Analysis*, Krasnoyarsk, Academy of Sciences Computing Center, Publication No. 17, 1990, pp. 35–40 (in Russian).

[76] V. Stoltenberg-Hansen, I. Lindström, and E. R. Griffor, *Mathematical Theory of Domains*, Cambridge University Press, Cambridge, 1994.

[77] A. N. Tikhonov and V. Y. Arsenin, *Solutions of Ill-Posaed Problems*, W. H. Whinston & Sons, Washington, D.C., 1977.

[78] A. S. Troelstra, *Lectures on Linear Logic*, CSLI, Stanford, California, 1992.

[79] G. S. Utyubaev, "On the ellipsoid method for a system of linear differential equations", In: S. P. Shary (ed.), *Interval Analysis*, Krasnoyarsk, Academy of Sciences Computing Center, Publication No. 16, 1990, pp. 29–32 (in Russian).

[80] J. Vehi, J. Rodellar, M. Sainz, and J. Armengol, "Analysis of the Robustness of Predictive Controllers via Modal Intervals", *Reliable Computing*, 2000, Vol. 6, No. 3, pp. 281–301.

[81] K. Villaverde and V. Kreinovich, "A linear-time algorithm that locates local extrema of a function of one variable from interval measurement results", *Interval Computations*, 1993, Vol. 4, pp. 176–194.

[82] K. Villaverde and V. Kreinovich, "Parallel algorithm that locates local extrema of a function of one variable from interval measurement results", *Reliable Computing*, 1995, Supplement (Extended Abstracts of APIC'95: International Workshop on Applications of Interval Computations, El Paso, Texas, Febr. 23–25, 1995), pp. 212–219.

[83] G. W. Walster, E. R. Hansen, and J. D. Pryce, *Practical, exception-free interval arithmetic on the extended reals*, Sun Microsystems Technical Report, 2001.

[84] G.-Q. Zhang, *Logic of Domains*, Birkhäuser, Boston, 1991.