Back to Contents
Next to Chapter 5

Chapter 4

Why Do We Use Probabilistic Concepts to Describe the World?[1]

Introduction

For a long time science had been developing so that scholars were satisfied with its results if they led to explanations of Nature or to its control (as in the natural sciences) or to inwardly consistent results (as was the case with mathematics). However, criticism began to arise as a result of the drive for self-analysis. Mathematics was the first scientific subject which already at the end of the last century manifested the need to construct its own foundations. The landmarks along this course are the discovery of contradictions in the theory of sets by Russell, Hilbert's intention to prove the absolute consistency of mathematical structures, ending in Gödel's theorem on undecidability, and, finally, metamathematics in its present-day state.

In the natural sciences the necessity to create their foundations was realized much later—this seems to have happened no earlier than in the 1930's. However, now we are able to speak about a fairly distinct and solid trend called "the philosophy of science."

Research into the logic of the development of science has not brought such powerful results as were obtained in metamathematics, but the papers by Carnap, Reichenbach, Popper, Kuhn, and Feyerabend are very interesting. In any case, they make us think about what we are doing in science. At the closing session of the Fifteenth International Congress of Philosophy in Varna (1973), we heard roughly the same: ". . . certain-

[1] This chapter was published in Russian as a preprint, "Language of Probabilistic Notions," by the Scientific Council of Cybernetics, Moscow, 1976. The major part of it was published in the journal *Automatika* (no. 1, 1979), which is translated into English. This chapter was translated by A. V. Yarkho.

Use of Probabilistic Concepts in Descriptions of the World

ly, philosophy has not moved humanity closer to the truth, but it has made it now more difficult to make mistakes."

One of the profound problems of the logic of science is why we use probabilistic concepts to describe the world, in contrast to the traditional deterministic mode of description. Have we any grounds to do this, and if so, what are they? I felt a desire to discuss these questions after reading the book *Theories of Probability. An Examination of Foundations* by Fine (1973) and the critical remarks on it by Tutubalin (1974).

Criticism by T. R. Fine and V. N. Tutubalin

The book by Fine is immediately interesting as a result of his use in the title of *theories*, rather than theory, of probability. Breaking the existing paradigm, he made an attempt to give a mutual comparison of all or almost all existing conceptions of probability, excluding only the rather incomprehensible fiducial probabilities of Fisher, non-commutative[2]

[2] This not widely known trend is developed in connection with certain problems of quantum mechanics, where we have to assume the existence of random values with deliberately large variance. For details, see the article by Parthasarathy (1970).

probability theory, and probabilistic concepts as model constructs in human language. The resulting book contains: the classical theory of Laplace, the frequency conceptions of von Mises, the frequency theory of Reichenbach–Salmon, Kolmogorov's axiomatics, the axiomatics of comparative probabilities (Fine's field of interest),[3] the algorithmic approach of Chaitin, Kolmogorov, and Solomonoff to evaluating randomness as complexity, the logical probability of Carnap, and the subjective (or personal) probabilities of Savage and Finetti. The variety of probability theories does not, strictly speaking, form a unified mathematical structure (in terms of Bourbaki). Rather, we observe a mosaic of separate logical structures, which have proved to be interrelated, to some extent. The author has managed to illustrate these interrelations explicitly. More than that, he has attempted, and not unsuccessfully, to demonstrate that each of these theories reflects essential features that we connect with the concept of probability. That leaves us small hope that a general all-embracing theory of probability can be constructed now. At the same time, while reading the book, one cannot help asking why the present-day tradition does not reflect the variety of probabilistic concepts in the textbooks? Even the algorithmic approach to a definition of randomness and probability is lacking. I do not really believe that many specialists in probability theory ever took note of the fact that logical probability proves compatible with frequency-based probability and, moreover, seems to be of great help in elucidating the classical probability concept.

However, I have to acknowledge that if a course of lectures embracing all the theories of probability were presented, the beauty of structures would immediately perish: the architecture of the course would become strange for a mathematican.

Still, the burden of Fine's book is not an attempt to elucidate connections among separate theories of probability, but rather to answer the question of why we use probabilistic concepts to describe the world. Do any theories of probability make this practice legitimate? Do theories of probability possess sufficient grounds for this? Fine answers this question in the negative. Here are some excerpts from his concluding remarks:

> *Finite Relative-Frequency Interpretation of Probability.* The advantage of a finite relative-frequency interpretation of probability is that it easily answers the measurement question, at least for those random phenomena that are unlinkedly repeatable. . . . Finite-relative frequency is descriptive of the past behavior of an experi-

[3] This is another trend that is not broadly known. It deals with the approach when the probability of an event does not take a numerical value but we can say about certain pairs of events that one of them is more probable than the other.

ment but it is difficult to justify as being predictive of future behavior. . . . Furthermore, not all random phenomena are amenable to analysis in terms of arbitrary repetitions. Why are not unique occurrences fit for probabilistic analysis? If we look closely, we see that we never exactly repeat any experiment. Nevertheless, it appears that informal recognitions of approximate repetitions coupled with a finite relative-frequency interpretation is the most commonly applied theory of probability.

Limit Relative-Frequency Interpretation. A limit statement without rates of convergence is an idealization that is unlike most of the idealizations in science. . . . Knowing the value of the limit without knowing how it is approached does not assist us in arriving at inferences. The relative-frequency-based theories are inadequate characterizations of chance.

Algorithmic Theory. The theory of computational-complexity-based probability, . . . while successful at categorizing sequences, is as yet insufficiently developed with respect to the concept of probability. However, the indications are that when developed, this approach will be able to measure probability, but will encounter difficulties with the justification of the use of the measured probabilities. The justification problem seems to be very similar to that faced by other logical theories of probability.

Classical Probability. Classical probability . . . is ambiguous as to the grounds for and methods of assessment of probability. It partakes of elements of the logical and subjective concepts and is far less clear than the logical theory as to how to reach probability assessments. It is also perhaps true that the subjectivist claim to subsume classical probability as a special case is valid. In the absence of a clear interpretation of classical probability, we cannot arrive at a determination of a justifiable role for it. . . . The axiomatic reformulations remove some of the measurement ambiguities but do little to advance the problem of justification.

Logical Probability. Formal processing of empirical statements need not lead to empirically valid conclusions. . . . Carnap . . . attempted to justify logical probability as being valuable for decision-making, but good decision-making requires more than just coherence, if it even requires that. Hence the measurement of logical probability and the justification of an application of the theory are as yet unsolved; the former appears more likely to be settled than the latter.

Subjective Probability. Of all the theories we have considered, subjective probability holds the best position with respect to the values of probability conclusions, however arrived at. . . . Unfortunately, the measurement problem in subjective probability is sizable and conceivably insurmountable. . . . The conflict between human capabilities and the norms of subjective probability often makes the measurement of subjective probability very difficult. (pp. 238–239)

Trying to explain the outward success of the probabilistic approach in physics, Fine says that these results are

1) irrelevant to inference and decision-making,
2) assured by unstated methodological practices of censoring data and selectively applying arguments,
3) a result of extraordinary good fortune. (p. 245–246)

And now a few words of Tutubalin's criticism. His constantly repeated criticism (see, e.g., Tutubalin, 1972) is not so depressing as that by Fine. However, reviewing the book by Fine, he writes:

> By now rather a spicy situation has formed, when many popular textbooks, following the tradition which dates back to *Analytical Theory of Probabilities* by Laplace, greatly exaggerate the significance and sphere of application of probability theory. . . . far from all such alluring achievements, its declarations are fraught with significant negative scientific consequences, simply because the authors of textbooks do not guide themselves by their declarations in concrete actions; . . . But if everything usually written in textbooks is first earnestly accepted and then critically analyzed, the results will still be discouraging.

As a matter of fact we are facing an odd situation. On the one hand, a broad and, it seems, fruitful development of statistical methods is going on – probabilistic thinking and its effect upon scientists' views are being discussed. On the other hand, we hear disappointing warnings from some mathematicians. How can this be accounted for?

I believe that theory – or, better, theories – of probabilities promoted the creation of a probabilistic language. The proper mathematical constituent of these theories is their mathematical *structure*, *grammars* of dialects of the probabilistic language. And for this reason the problem itself of logical foundations on which to base the legitimacy of applying a specific probability theory seems quite meaningless. It is more fruitful to speak of using the probabilistic language to describe phenomena of the external world, this language being significantly softer than the traditional one, based on causal relations. Generally speaking, language can be convenient or inconvenient to describe something. The legitimacy of what we say in any language is given not by the structure (grammar) of the language but by the way we support our statements. Grammar serves only to make the phrases correct, i.e., corresponding to the rules of inference. Language is certain to influence the peculiarities of our argumentation. On some occasions it requires rigid causal relations, and on others it allows us to confine ourselves to vague but, somehow, normed judgments. A conversation without *any* rules seems to lack sense altogether.

We will start to develop our idea of a probabilistic language from the history of familiar deterministic conceptions of the world and the role this system ascribed to chance.

History of Determinism

Determinism as a concept has two meanings. Broadly, it is an unconditional belief in the power and omnipotence of formal logic as an instrument for cognition and description of the external world. In the narrow sense, it is the belief that all phenomena and events of the world obey causal laws. Furthermore, it implies confidence in the possibility of discovering, at least in principle, those laws to which world cognition is reduced.

The causal interpretation of the phenomena of the external world seems to be characteristic of the earliest forms of human thinking. At least, primitive tribes observable at present, with their rather alien forms of pre-logical thinking, lack the notion of chance altogether. To them, everything is mutually interrelated and predestined; all phenomena are perceived as signs or symbols of something. Levy-Brühl (1931), a well-known ethnologist of the recent past, describes this system of ideas:

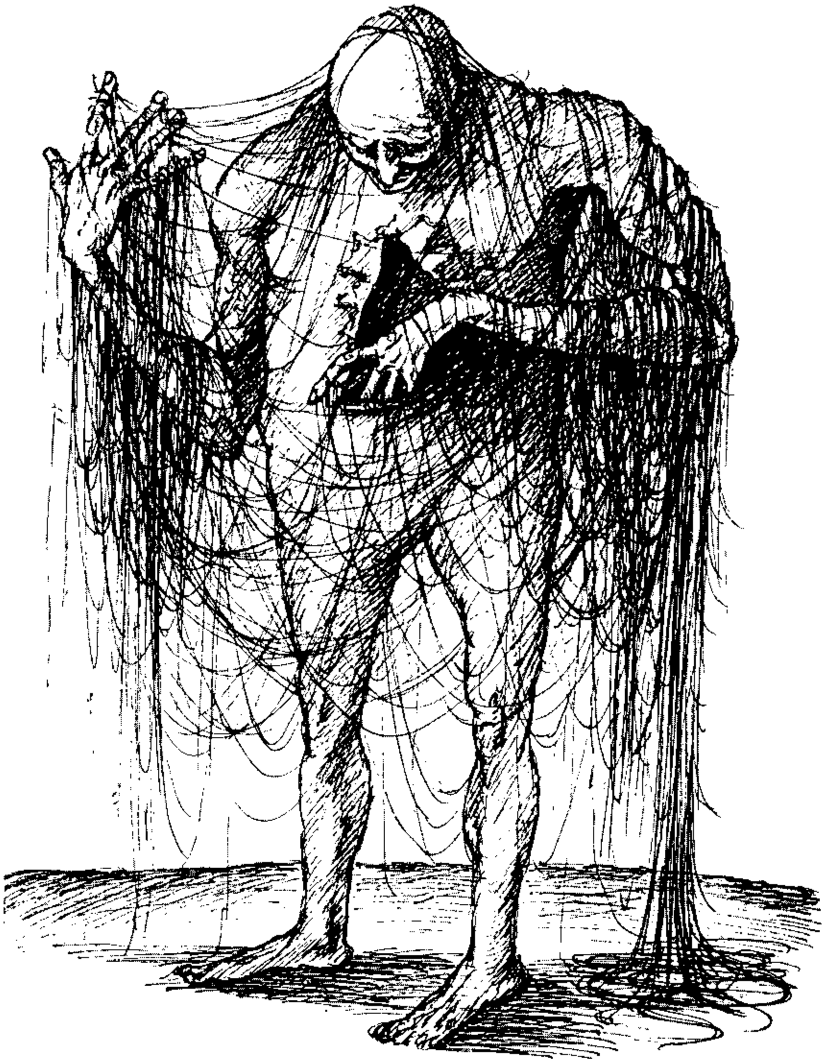
> . . . For spirits so disposed, there is no chance; they do not overlook what we call the fortuitous. But as for a true accident, a sorrow, small or great, is never insignificant, it is for them always a revelation, a symbol, and has its reason in an invisible power which thus manifests itself. Far from being due to chance, it itself reveals its cause.

If these observations are extrapolated into the past, it allows us to believe that humanity at its early stages perceived and described the world in the language of irrational or even mystic causal notions.

The well-known ancient Indian teaching of Karma, as a large-scale system rigidly and meticulously determining man's fate through his actions in previous lives, may be regarded as a concise remnant of a once vague conception of the universal *causal* basis of the world.

Unfortunately, we do not possess the data which would allow us to trace the whole complicated progress of human thought which transformed ambiguous mystic ideas of causal relations into the logical structures of European thought. It is important to note that Aristotle classified and codified logical forms of expressing thought.[4] Later, in the Middle Ages, almost all intellectual life was devoted to the attempt to comprehend the role of logic in the universe and in human thinking. The

[4] Here is how Aristotle estimated the cognitive role of maintaining causal relations (quoted from Akhmanov, 1960, p. 159): "We believe we know every thing in a simple way when we believe we know the cause of its existence, and know not only its being the cause, but also that it cannot be otherwise."



Cause-Effect Links

medieval scholastics are responsible for the development and strengthening of rigid determinism.[5] They popularized (as well as vulgarized) Aristotle and made the study of logic universal. A very significant role was

[5] It may be supposed that determinism is deeply rooted in the widely known Cabala of the Middle Ages; this peculiar algebra of belief was built as a calculus over 22 letters-symbols, each of which corresponded to names and numbers. But who can trace the way this secret teaching influenced European thinking? However, many important analogies do arise, e.g., with the universal symbolism of Leibniz. Interest in the Cabala in the Middle Ages seems to have stimulated the construction of logical automata, of which so many beautiful tales are told.

played by Thomas Aquinas, the founder of Thomism, which was one of the two dominant trends of scholasticism. Here is one description of his outlook (Dragunov, 1970):

> Thomas defined the truth as "adequate correspondence (adequatio) between mind and thing . . ." (per se notum) and an unconditional principle of thinking and being, as well as a criterion for truly rational cognition. (p. 381)

The reader is here facing a very strong, though fairly naive, postulate of consistency both for thinking and for the processes of the world. Such axiomatics makes cognition and description of the world much easier, especially if the first postulate is accepted and interpreted as a primitive version of the reflection theory of truth accepted by Marxism.

The following words by William of Occam, a scholastic of the early fourteenth century who represented late nominalism, are also of interest (Styazhkin, 1967):

> Logic, rhetoric and grammar are not speculative subjects but genuinely cognitive guides since they really govern the mind in its activity. (p. 143)

The role of Thomas Aquinas in the evolution of European thinking was fixed by Pope Leo XIII in his encyclical "Acterni partis" in 1879, where Aquinas's philosophic-theological system was acknowledged as "the only true philosophy of catholicism" (Subbotin, 1972). It is probable that this papal intervention also contributed to the deification of a rigid determinism.[6]

An important role in understanding causality was played by Kant. According to him, space, time, and causality are prior forms of pure intellect, inherent categories which make our experience possible. Was this not the insight that causal and space–time arrangements of the observed phenomena are precisely a result of our language?

I next quote a very sharp statement concerning space, time, and causality made by A. D. Aleksandrov, a well-known contemporary physicist, in a paper devoted to philosophical comprehension of relativity theory (Aleksandrov, 1973):

[6] It is noteworthy that unofficial European religious–philosophical thinking, including the esoteric schools, is based upon the belief in rigid determinism. For example, in the beginning of the twentieth century there appeared the exposition of ostensibly ancient Egyptian teaching (previously concealed) where determinism was proclaimed, among other principles (Stranden, 1914, pp. 72–73): "Every cause has its effects; every effect has its cause; everything goes according to the law; chance is only a name by which we call the laws unknown to us; there are many aspects of causality, but nothing escapes the law." It is hardly wise to try to discover the true age of these statements. But, judging by their context, we may assume them to be a modernized and Europeanized exposition of some ancient oral concepts. I have quoted these words, unusual for a scientific paper, only to show to what extent determinism had penetrated European thinking.

> The space-time structure of the world is nothing but its causal structure, but taken in a correspondingly abstract form. This abstraction consists in omitting all the properties of phenomena and their causal relations, except those indicating that phenomena are made of events and their mutual effects are accumulated from the influence of some events upon others.

This quotation, certainly, is not a direct return to the Kantian notion of the inherent nature of space and time, since the notion of inherence here has a different—modern—sense: inherent is the capacity for profound abstraction. It is easy to assume that this capacity developed in the process of evolution and became genetically fixed, and in this sense (and not in some metaphysical-idealistic one) it became inherent. But the most important feature here is that if, in dealing with notions of space, time, and causality, we have to acknowledge the existence of certain universal structures of our perception, which arise as a result of abstraction, we thus acknowledge their linguistic nature. Without these structures of perception we could not have discussed our observations of the external events. This immediately gives rise to the question: Are we dealing with the only possible system of structuring or are other linguistic categories also possible?

Now let us consider the role of determinism in the development of science. It began with classical mechanics, the simplest theory based upon determinism (at least so it seems at first glance). From school years we are brought up with the idea that by applying the laws of classical mechanics it is possible to predict the future of a material system, its initial data being known. Later, we reach the conclusion that classical mechanics is the best example of our knowledge.

As a matter of fact, everything we obtain by means of classical *mechanics is nothing more than an approximate description.* Strictly speaking, initial data are never known with certainty; the only information which may be evident is their distribution. Further, during motion unpredicted random forces can influence a system; at least, it is not likely that a given system will remain isolated during the period in which we are going to make predictions. Even such an accurate branch of science as celestial mechanics needs corrections from time to time. All this is well analyzed in a highly readable book by Blokhintsev (1966). I shall only remark that classical mechanics has holes in a purely theoretical sense. Bohr (1955) drew attention to the fact that the law of inertia violates the principle of causality: a uniformly moving body in a vacuum keeps moving without any cause.

However, when the laws of classical mechanics were applied, the approximation often proved so accurate that it struck scientists as a miracle. This high degree of accuracy was accounted for by the fact that people succeeded in selecting phenomena invariant in relation to the system in which they take place. Wigner (1960) believes that Galileo must have been puzzled by the fact that stones fall from a tower in a manner independent of their size, the weather, or who throws them. Classical physics, too, studied the phenomena invariant in relation to the changing states of the system. One of the favorite questions of today asks whether chemistry or even biology is reducible to physics. But it turns out that not even all physical problems can be reduced to a description in the framework of traditional physical concepts.

This is true, first of all, of problems of technical physics, where the object of research is the system itself. One of my favorite examples of such systems is sprectrochemical analysis. Here we are dealing with the system in which many well-known physical phenomena act simultaneously: hydrodynamic flow by discharge, explosion evaporation, equilibrium evaporation, selective oxidation on the electrodes, diffusion in a solid (influenced by the solid's structure), emergence of a gas cloud, diffusion within it, excitation of atoms and radiation—and on top of all this, inaccurate sharpening and installing of electrodes and unstable parameters of the excitation generator. In this system, invariants—dominant phenomena—cannot be selected. Nothing can be described in familiar terms of physics, though it is always possible to think of an experiment in which nearly every one of the enumerated phenomena could be considered almost as an invariant. However, the problem is formulated so as to enable the study of the whole system.

True, still earlier physicists had to face the impossibility of selecting dynamic invariants in constructing the kinetic theory of gases; in order to connect molecular processes with the macroscopic state of a system, they had to introduce a probabilistic description.

Later, in quantum mechanics, the change of scales made the concept of the precise particle location in space impossible, whence comes the impossibility of the familiar notions of phenomena arranged in time and space.

But all this is known only too well. I remind the reader of the evolution of thinking in physics only in order to answer the question analogous to that asked by Fine in discussing the legitimacy of probabilistic conceptions for description of the external world. The question can be stated as follows: Have we sufficient logical grounds to describe the external world with deterministic concepts? These grounds, if any, are more of a historical-psychological nature than a logical one. Human prehistory prepared people for causal interpretation of phenomena. Medieval scholastics strengthened and deepened belief in "determinism" in its broad meaning. Progress in classical mechanics fixed the belief into the causal picture of the world for a long time. Later scientific development, especially in physics, continued to contribute to this belief, but eventually it began to shake it loose.

If we wish to justify rigid determinism in its broad sense from a strictly logical position, we have to recognize as axioms Thomas Aquinas's statements that thinking is consistent, the world is consistent, and consistency is the criterion of the truth. Hilbert's program, directed at proving the absolute consistency of mathematical structures, and the program of the neopositivists in the form in which it was set forth by Carnap—are these not a distant echo of the postulates of Thomas Aquinas? After both these programs failed, and especially after the appearance of Gödel's theorem on undecidability and the progress of quantum mechanics, what can we say in favor of the absolute belief in determinism? Although I am quite aware of the fact that certain outstanding scholars of the recent past, say, Einstein, were determinists (his discussion with Bohr on the subject is well known), it seems more pertinent to speak here of the paradigm of the epoch rather than of clearly formulated logical foundations.

History of the Teaching of Chance

I have already pointed out that the concept of chance was quite foreign to the psychology of primitive people. It is impossible to trace in any detail the history of the emergence and formulation of such concepts as probability and chance. Only scanty information is available.

The intellectually rich society of Ancient India lacked the concept of probability in its modern meaning, though Indian thinkers understood only too well the universal changeability of the world and approached rather closely contemporary ideas of stochastic processes. This can be illustrated by the famous dialogue between Milinda and Nagasena (Oldenberg, 1881). This dialogue is a fragment of an historical document which records the account of the meetings between Menander (Milinda), a Greek prince who ruled on the territory of the Indus and in the valley of the Ganges in 125–195 B.C., and Nagasena, a Buddhist teacher.

> "It is as if, sire, some person might light a lamp. Would it burn all night long?"
>
> "Yes, revered sir, it might burn all night long."
>
> "Is the flame of the first watch the same as the flame of the middle watch?"
>
> "No, revered sir."
>
> "Is the flame of the middle watch the same as the flame of the third watch?"
>
> "No, revered sir."
>
> "Is it then, sire, that the lamp in the first watch was one thing, the

lamp in the middle watch another, and the lamp in the last watch still another?'

"O, no, revered sir, it was burning all through the night in dependence on itself."

"Even so, sire, a continuity of dhammas runs on, one uprises, another ceases; it runs on as though there were no before, no after; consequently neither the one (dhamma) nor another is reckoned as the last consciousness."

What can be said of the causal arrangement of phenomena in time and space if such an outlook is shared? Does this not resemble a modern description of a random process?

The well-known Indian statistician Mahalanobis tried to trace a certain analogy between modern statistical theory and the ideas of Jainist logic [the religion and philosophy of Jaina, which reached its full blossom in the times of Great Mahavira (589–527 B.C.), Buddha's contemporary]. Jainism contained a system of ideas called syādvada, close to the modern probabilistic concepts (Mahalanobis, 1954).

The *syādvada* is set forth as follows: (1) May be, it is; (2) may be, it is not; (3) may be, it is and it is not; (4) may be, it is indescribable; (5) may be, it is and yet is indescribable; (6) may be, it is not and it is also indescribable; (7) may be, it is and it is not and it is also indescribable.

. . . all things are related in one way or the other and . . . relations induce relational qualities in the relata, which accordingly become infinitely diversified at each moment and throughout their career. . . . Things are neither momentary nor uniform.

A reality is that which not only originates, but is also liable to cease and at the same time is capable of persisting. Existence, cessation, and persistence are the fundamental characteristics of all that is real. . . . This concept of reality is the only one which can avoid the conclusion that the world of plurality is the world of experience, is an illusion. (p. 103)

One cannot help wondering, while reading these ancient fragments, at how far they had progressed as compared with primitive notions of simple space–time arrangement. And how far is the conception of syādvada from the limitations imposed upon thinking by the laws of formal logic. Still, all this coexisted with the universally accepted notion of karma—a system most rigidly arranged on a large time scale.

It would be very interesting to trace the pre-history of probabilistic notions in European thinking, but this is very difficult to do. The point is that neither the Hellenic epoch nor the Middle Ages gave birth to any coherent probabilistic conceptions. There occurred only separate and often contradictory statements about the role of chance, which could have many interpretations. Still, I shall try to show how elements of probabilistic judgments began to ooze through a general deterministic background. I base my argument principally upon the very interesting and thorough paper by Sheynin (1974).

It is difficult to say anything definite about the way the role of chance was estimated by Greek atomists. They were strict determinists. Democritus rejected chance quite obviously. Well known is the statement by Lysippus: "Nothing comes from nothing but everything comes from foundations and necessity." At the same time, the atomists were reproached by their contemporaries because they would ascribe everything to chance since this logically followed from their constructions [for more details, see Russell (1962, p. 85) and Sheynin (1974, p. 102)].

Chance is repeatedly mentioned by Aristotle. According to Sheynin, it was Aristotle who introduced the concept of chance and accident into classical philosophy, defining it as follows: accident "is something which may possibly either belong or not belong to any one and the selfsame thing . . ." (Sheynin, 1974, p. 98). Sheynin draws our attention to the fact that Aristotle's works also contain reasoning on the probable. A probability, says Aristotle,

> is a generally approved proposition: what men know to happen or not to happen, to be or not to be, for the most part thus and thus . . ., e.g. "the envious hate" . . . (Sheynin, 1974, p. 101)

Below are several more statements by Aristotle, which I cite as they are given by Sheynin (1974):

> As to chance (and change) they are "characteristic of the perishable things of the earth" . . . Some effects could be caused incidentally, i.e. by spontaneity and chance, chance is opposed to mind and reason and its cause "cannot be determined. The products of chance and fortune are opposed to what is, or comes to be, always or usually." (p. 98)

The general impression is that Aristotle, acknowledging the role of chance in life, attributed it to something which violates order and remains beyond one's scope of comprehension. He did not recognize the possibility of a science of chance, though he understood that various human activities are connected with it. Aristotle said that in navigation "not the cleverest are the most fortunate, but it is as in throwing dice" (p. 101). He described rhetoric as an art of persuasion based on probabilities. As Sheynin (1974) points out, he even introduced a rudimentary scale of subjective probabilities stating that "a likely impossibility is always preferable to an unconvincing possibility." (p. 101).

Now let us see what Thomas Aquinas's attitude toward chance was.

Here are several statements from his famous tractatus *Summa Theologica* [again cited from Sheynin's article (1974)]:

> (1) The effects willed by God happen contingently . . . because God has prepared contingent causes for them.
> (2) Casual and chance events are such as proceed from their causes in the minority of cases and are quite unknown. (p. 103)
>
> . . . some causes are so ordered to their effects as to produce them not of necessity but in the majority of cases, and in the minority to fail in producing them . . . which is due to some hindering cause . . . (p. 103)
>
> . . . if we consider the objects of science in their universal principles, then all science is of necessary things. But if we consider the things themselves, then some sciences are of necessary things, some of contingent things. (p. 104)

The conception of chance shared by Aquinas is, of course, hard to outline very clearly. Sheynin remarks that, according to Byrne (1968), there is something in common between Aquinas's conception of probability and modern logical probability theory as well as between his theory of contingency and the modern frequency theory. However, be that as it may, Aquinas's views on the nature of chance did not influence the progress of probabilistic concepts in modern times. The quotations cited above are interesting for us in that they illuminate the way an outstanding thinker of his times, an ardent believer in logic, tried to cope with chance.

Medieval scholastics also faced the necessity to comprehend another concept in the framework of logical notions: "free will." Logically speaking, the concept of free will is equivalent to that of chance. If in an experiment with tossing a coin many times, we assume that in each given fall, in accordance with the concept of chance, the coin lands unpredictably, this is logically equivalent to ascribing to it free will with certain well-known statistical limitations laid upon the set of tosses of the coin. Free behavior is as difficult to ground logically as the possibility of chance is. Not without reason is the problem of free will one of the "accursed" questions of philosophy.[7] Already Buridan, a French scholastic, believed the problem of free will to be logically unsolvable. His argu-

[7] Attempts to find logical grounds for the concept of freedom have been made in our times, too. This is illustrated by an elegant paper by Gill (1971), where freedom is considered in the framework of a calculus—in order to define freedom, three postulates are introduced, and two theorems are formulated. In constructing a calculus, chance, naturally, has to be excluded. This is formulated in the following manner: "If a given command is contingent, its contradictory opinion is necessary—it cannot be rejected without self-contradiction. The agent cannot control himself if he commands or permits an inconsistency. If a prohibition is contingent—not necessary—its contradictory permission is necessary" (p. 9). This is a brilliant sample of strictly formal reasoning in the style of logical positivism, showing that chance makes a system internally inconsistent.

ments seem to have influenced the later development of the problem. In any case, a well-known paradox is that of the *pons asinorum* (falsely ascribed to Buridan) about the donkey starving to death between two bales of hay as a result of the absence of logical grounds for decision making. The possibility of a random choice is here excluded as illogical.

Among modern philosophers, many outstanding thinkers either did not acknowledge the role of chance, or if they did, they connected it with numerous unknown causes. For example, the Dutch philosopher B. Spinoza (1632–1677) wrote:

> *Prop. XXIX.* Nothing in the universe is contingent, but all things are conditioned to exist and operate in a particular manner by the necessity of the divine nature.
> *Proof.* Whatsoever is, is in God, for he exists necessarily, and not contingently. Further, the modes of the divine nature follow therefrom necessarily, and not contingently . . . (Spinoza, 1955)

The British philosopher Thomas Hobbes wrote:

> (1) . . . generally all contingents have their necessary causes . . . but are called contingent in respect of other events upon which they do not depend . . .
> (2) . . . by contingent, men . . . mean . . . that which hath not for cause anything that we perceive . . . (cited from Sheynin, 1974)

The French philosopher C. A. Helvetius (1715–1771), in his famous tractatus *On Mind*, wrote:

> . . . chance; that is, an infinite number of events, with respect to which our *ignorance will not permit us to perceive their causes, and* the chain that connects them together. Now, this chance has a greater share in our education than is imagined. It is this that places certain objects before us and, in consequence of this, occasions more happy ideas, and sometimes leads to the greatest discoveries. . . . If chance be generally acknowledged to be the author of most discoveries in almost all the arts, and if in speculative sciences its power be less sensibly perceived, it is not perhaps less real . . . (Helvetius, 1809, p. 221)

According to the French philosopher P. H. T. Holbach (1723–1789):

> . . . Chance, a word devoid of sense , which we always oppose to intelligence without coupling it with a clear idea. In fact, we attribute to chance all those effects concerning which we see no link with their causes. Thus, we use the word chance to cover our ignorance of the natural causes which produce the effects that we see. These act by means that we have no idea of or they act in a manner in which we do not see any order or system; followed by actions similar to our own. Whenever we see, or believe we see order, we attribute this order to

> an intelligence, a quality derived from ourselves and from our fashion of acting and being affected. (Holbach, 1770)

We read in the works of the British philosopher David Hume (1711–1776):

> . . . chance is nothing real in itself, and, properly speaking, is merely the negation of a cause. . . . it produces a total indifference in the mind . . . the chances present all these sides of the die as equal, and make us consider every one of them, one after another, as alike probable and possible. . . . the chance or indifference lies only in our judgement on account of our imperfect knowledge, not in the things themselves, which are in every case equally necessary. (Hume, 1964)

Immanuel Kant wrote:

> In a body these absurdities were taken to such an extreme that they ascribed the origin of all living creation precisely to this blind concourse and actually derived reason from unreason. In my own concept, on the other hand, I find matter/substance bound to certain, distinct, necessary laws. . . . there exists a System of all Systems, a limitless understanding, and an independent Wisdom from which Nature also derives her origins according to her possibilities in the entire sum of determinations. (Kant, 1912)

Kant's *Critique of Pure Reason* contains a statement already acknowledging chance, at least in its individual manifestation: ". . . the individual accident (chance) is nevertheless entirely subordinated to a principle (rule)." This is, if you like, a concession to chance made within the boundaries of determinism. The German physician, naturalist, and philosopher L. Büchner (1824–1899) wrote:

> What we call chance is exclusively founded upon the tangle of circumstances, whose inner relations and final causes we cannot discover. (Büchner, 1891)

Hegel made a much more profound attempt to comprehend chance. In his *Science of Logic* we find the following statements:

> This union of Possibility and Actuality is Contingency. [The Contingent] has no foundation. The Contingent is indeed Reality as only that which is possible . . . this has a foundation.
>
> The contingent therefore, in consequence, because it is accidental has no ground, but even so it has a ground just because it is accidental. Here the union of necessity and contingency is itself present; this union is called Absolute Reality. (Hegel, 1971)

The words of Hegel, as usual, cannot be comprehended completely. But it is important to note that they lack naive negligence of chance and its reduction to uncomprehended or undiscovered causes. Engels wrote:

> . . . where on the surface accident holds sway, there actually it is always governed by inner hidden laws and it is only a matter of discovering these laws. (Engels, 1973, p. 48)

There was a time when Soviet philosophical literature displayed an acutely hostile attitude toward chance as a philosophical category. For example, *Concise Dictionary of Philosophy* edited in 1955 contained the following opinion (Rosental and Yudin, 1955):

> Cognition may be considered scientific only so far as it acknowledges the natural and social phenomena in their necessity. Cognition cannot be based on randomness. Behind randomness science always strives to discover regularity and necessity (pp. 325–326).

Later, such extreme judgments were recognized as erroneous. Now the following statement, given in the *Philosophical Encyclopaedia* published in 1970, seems to be considered correct:

> Science by no means stops at randomness but strives to understand regularity and necessity. But recognizing the objectivity of randomness, we have to recognize the necessity to study it. Random phenomena and processes are a special object of several modern sciences, including physics, biology, sociology, etc. Such branches of modern mathematics as theory of probability, theory of random functions, theory of stochastic processes, are completely devoted to studying quantitative characteristics of chance (Yakhot, 1970, p. 34).

This text is already a significant step forward. Science gets the right to study chance, though it is said that it does not stop there, but strives to understand regularity. It only remains unclear how to pass from quantitative parameters to understanding necessity.

The naive belief that chance emerges in our consciousness as a consequence of ignorance passed from philosophy to the natural sciences, where it was shared by Galileo, Kepler, Huygens, Bernoulli, Lambert, and even Laplace. Here is what Kepler said of chance:

> But what is Chance? Nothing but an idol, and the most detestable of idols—nothing but contempt of the sovereign and all-powerful God as well as the very perfect world that came from his hands. (cited from Sheynin, 1974, p. 127)

And here are the words of Laplace, one of the creators of probability theory:

> Chance has no reality in-itself; it is nothing but the proper terms for designating our ignorance of the manner by which the different parts of a phenomenon coordinate among themselves and the rest of Nature. (cited from Sheynin, 1974, p. 132)

I would also like to mention the thoughts of two outstanding scholars of comparatively modern times on the subject of chance. Darwin in his *Origin of Species* wrote:

> I have hitherto sometimes spoken as if the variations . . . had been due to chance. This, of course, is a wholly incorrect expression, but it serves to acknowledge plainly our ignorance of the cause of each particular variation. (cited from Sheynin, 1974, p. 115)

Henri Poincaré, one of the most prominent mathematicians of the recent past and one of the first scholars interested in the philosophical foundations of science, also believed [as pointed out by Sheynin (1974)] that chance has an influence when, under the conditions of unstable equilibrium, very weak causes produce a very strong effect (see Poincaré, 1952).

The concept of chance was initially introduced into science by physicists, at the end of the twentieth century. They seemed to feel quite unmoved by the problem of the philosophical comprehension of chance. They had to explain and describe the world, and this description did not fit the limits of deterministic conceptions. Certain phenomena could only be well described in probabilistic language. The landmarks of this process are well known: creation of kinetic theory of matter by Maxwell and Boltzmann; the latter's statement that our world is but a result of a huge fluctuation; introduction of the notion of an ensemble by Gibbs and the canonical distribution discovered by him (this led not only to the creation of statistical physics but to something more—to forming a new outlook in physics); the study of Brownian motion, which gave impetus to developing the theory of random functions; and, at last, the progress of quantum mechanics. But they were not worried about the philosophical problem or logical foundations of the legitimacy of this approach. The world of observed phenomena was well described, and this was a sufficient foundation. The sorrowful ponderings of the philosophers of the past about chance were merely forgotten. Here are the thoughts of the well-known physicist Max Born on the relation of randomness and determinism:

> We have seen how classical physics struggled in vain to reconcile growing quantitative observations with preconceived ideas on causality, derived from everyday experience but raised to the level of metaphysical postulates, and how it fought a losing battle against the intrusion of chance. Today the order of ideas has been reversed: chance has become the primary notion, mechanics an expression of its quantitative laws, and the overwhelming evidence of causality with all its attributes in the realm of ordinary experience is satisfactorily explained by the statistical laws of large numbers. (Born, 1949, p. 120–121)
>
> . . . I think chance is a more fundamental conception than causality;

for whether in a concrete case a cause-effect relation holds or not can only be judged by applying the laws of chance to the observations (Born 1949, p. 47)

However, this is only a panegyric to chance; in no way is it a logical analysis of what chance is.

I shall not here dwell on the development of probabilistic concepts in mathematics. The early period—the eighteenth century and the beginning of the nineteenth century—is thoroughly illuminated in the papers by Sheynin (1971*a*, *b*, 1972*a*, *b*, 1973*a*, *b*, *c*). The later period is well known to everybody who is interested in probabilistic concepts. I shall only make one brief remark. When probabilistic methods in mathematics began to develop, they proceeded not from some general concepts of the insufficiency of deterministic methods to describe the phenomena of the external World (analogous, say, to the philosophy of Jainism) but from the attempt to describe and comprehend two quite particular phenomena: on the one hand games of chance and on the other hand elaboration of the theory of errors which resulted from the introduction of degree measurements in the instrumental astronomy in the seventeenth and eighteenth centuries (for details, see the papers by Sheynin).

Mathematical statistics in its modern form was created only at the end of the nineteenth and the beginning of the twentieth century, after the publication of papers by F. Galton, K. Pearson, and R. Fisher. Then probabilistic methods of research began to penetrate into various fields of knowledge.

Formation of a Probabilistic Paradigm

The theory of probability or, better, theories of probabilities of the present create something more than a theory for describing mass, repeated phenomena: they generate a new paradigm that allows one to describe the observed world in a weaker language than that of the rigid deterministic ideas traditionally accepted in science.

Language of probabilistic concepts. I shall try to elucidate this idea in detail. We say that a random value is given if its distribution function is given. That means that we quite consciously abandon the causal interpretation of the observed phenomena. We are satisfied with a purely *behavioral* description of phenomena. A distribution function is a description of random value behavior, without any appeal to what has caused this type of behavior. At last, we acquire the right to describe a phenomenon simply as it is. Moreover, the description is given in some blurred, uncertain way: the probability that a continuous random value *in its realization* (say, as a result of measuring) will occupy some fixed point equals zero. We can speak only of the probability that some random value will fall within an interval of values.

Does not this imply quite a novel view of the world or, at least, the possibility of a description radically different from the traditional deterministic one?

Let us examine the well-known illustration with tossing a coin. Remaining in the probabilistic position, we assume that in each separate tossing a coin may fall as it likes; i.e., as I have already said, we ascribe *to the coin free will, though we also lay statistical limitations upon the* result of large numbers of tests. This is a much weaker description of a phenomenon than an attempt to predetermine, proceeding from the laws of mechanics, in what way the coin will fall. At first sight, it seems that the chain of causal phenomena leading to a concrete result in a concrete act of tossing the coin may be traced in the main. But if we try to do this, we shall immediately have to introduce into consideration an incredibly large, perhaps infinitely large, number of facts and circumstances, and our chain of causal links will have to be extended to include space phenomena rooted in some immensely remote past that is unknown to us.[8] It is noteworthy that tossing a coin is almost the same as the experiment with which Galileo started the progress in mechanics. However, in one formulation of the problem, the experiment with throwing proves invariant to the surrounding phenomena, whereas in another formulation this is not so.

All that was said above pertains not only to tossing a coin or dice. It also pertains to the behavior of error in any experiment as well as to the behavior of any sufficiently complicated system. As mentioned above, Darwin thought an attempt to explain variation in biology by chance should not be taken seriously. However, at present we have every reason to believe that the origin of species cannot be regarded as a result of a rigidly given program. Mutations have to be connected with chance. This follows both from biological considerations (Monod, 1972) and from logical ones (Nalimov and Mul'chenko, 1970; see also Chapter 7 of this book). From Gödel's proof of undecidability, it clearly follows that any sufficiently rich logical system is incomplete, and extended, but finite,

[8] It is of interest to quote here statements by Max Born concerning the difficulty of understanding the idea of a "causal chain" (Born, 1949, p. 129): "One often finds the idea of a 'causal chain' A_1, A_2, . . . where B depends directly on A_1, A_1 on A_2, etc., so that B depends indirectly on any of the A_n. As the series may never end, where is a 'first cause' to be found?—the number of causes may be, and will be in general, infinite. But there seems to be not the slightest reason to assume only one such chain, or even a number of chains; for the causes may be interlocked in a complicated way, and a 'network' of causes (even in a multidimensional space) seems to be a more appropriate picture. Yet why should it be enumerable at all? The 'set of all causes' of an event seem to me a notion just as dangerous as the notions which lead to logical paradoxes of the type discovered by Russell. It is a metaphysical idea which has produced much futile controversy."

expansion of its axioms does not make it complete. In the language of such a system, true statements may be formulated which do not immediately follow from it, as well as false statements which will not be refuted. A deterministic description of the world as a whole, or even of a large subsystem such as the biosphere, must remain impossible. Yet, a consistent description of the world by appeal to chance seems intuitively possible.

The impossibility of accurately locating a particle, revealed in quantum mechanics, also mandates a "blurred" description of observed phenomena by means of probability waves. It is as a consequence of this weakened type of description that the causal nature of the system's progress can be preserved. In the words of Born (1949):

> [in quantum mechanics] we have the paradoxical situation that observable events obey laws of chance, but that the probability for these events itself spreads according to laws which are in all essential features causal laws. (p. 103)

The introduction of probability waves in quantum mechanics is, if you like, just the weakening of the rigid causal concepts of classical physics. The development of a wave is predictable during the observation, but prediction itself is of a non-deterministic nature to which we are accustomed in everyday life. The logic of reasoning is such that the causal progress of events is not completed. It breaks somewhere and is replaced by a probabilistic description of behavior.

An algorithmic definition of randomness as the complexity of a message can also be interpreted as a behavioral description. If we deal with a sequence of numbers consisting of zeros and ones, then, roughly speaking, complexity will be characterized by the minimal number of binary digits necessary to replace the sequence in transmitting it through a communication channel. According to A. N. Kolmogorov, those elements of a large finite aggregate of symbols are called random which have the greatest complexity. The concept of randomness emerges here from observing the behavior of a symbolic sequence. If it is impossible to discover an algorithm for generating numbers which would be simpler than the sequence, then the whole sequence must be transmitted through the communication channels. Such a sequence is naturally called random.

Fine (1973) tries to contrast determinism to chance in the following manner:

> We can distinguish between deterministic and chance phenomena capable of generating an indefinitely long sequence of discrete-valued outcomes on the grounds that deterministic phenomena yield outcomes of bounded complexity, whereas chance phenomena yield

> outcomes for which the complexity of increasing longer outcomes diverges. Probabilistic phenomena might then be characterized as the subset of chance phenomena for which the various outcomes have apparently convergent relative-frequencies. (p. 153)

Any algorithmic definition of a random sequence is clearly linguistic. Roughly speaking, we call random what we cannot describe briefly. And this is where language relativism immediately shows up. Imagine that we are dealing with numbers π and e. It is clear that there is no necessity to transmit through a communication channel all the figures giving the approximate value of these numbers: it will suffice to give the algorithm of their calculation. In this sense symbolic sequences approximately giving π and e are not random. At the same time it is known that these sequences of numbers are sometimes used as random ones in problems of simulation by the Monte-Carlo method. Indeed, statistical criteria we have at our disposal do not allow us to differentiate these sequences from those given by a meter registering radioactive decay. Now imagine that a symbolic sequence of π is recorded with the first symbols omitted. Who will guess that the sequence is not random? (True, this is not the only trouble.) The algorithmic approach is fraught with difficulties resulting from a particular choice of calculation programs. The conception, as a whole, is far from being complete.

Besides Kolmogorov's definition there are also definitions of randomness for infinite sequences, given by Donald W. Loveland, P. Martin-Lof, and G. J. Chaitin. Several definitions of probability based on the evaluation of complexity are proposed by R. J. Solomonoff. I have noted Kolmogorov's statements on this point. A more detailed discussion of the difficulties connected with the elaboration of algorithmic randomness is presented by Fine (1973), whose book also contains a substantial bibliography on the subject.

Axiomatics of the theory of probability as grammar. If probability theory in its applications is regarded as a language, its structure, given by the axiomatics, will be just the grammar of this language (Nalimov, 1974*a*; see also Chapter 3 of this book). By this approach we immediately avoid all Fine's (1973) lamentations that, from the foundations of probability theory, nothing follows concerning the possibility of its application. Any grammar, according to the meaning of the word, is aimed only at constructing grammatical and comprehensible—meaningful and consistent, or at least roughly consistent—phrases. But from grammar nothing ever follows concerning a language's applicability.

It may seem that the axiomatics of probability theory [we shall consider here only generally acknowledged axiomatics (Kolmogorov, 1956)] is, indeed, used as grammar; i.e., one has to fall back upon it while constructing comprehensible phrases. I would like to illustrate this statement by some examples. When a probabilistic statement is made, it is first of all necessary to be aware of the space of elementary events on which the probabilities are given. Otherwise, we can get absurd results such as probability greater than unity.[9]

The concept of σ-algebra gives a clear idea of the set of elementary events under consideration. One of the requirements here is: if A belongs to the set of events, then $\bar{A}$ (i.e., not A) also belongs to it, which is to say that a grammatically correct system of statements is built so that all possible logical operations remain within σ-algebra.

This is important if one is to understand texts containing probabilistic judgments.

Axioms of norming and non-negativeness are of great importance for understanding probabilistic statements. If we come across a statement containing negative probability, it will simply remain incomprehensible. And if we try to record undetermined behavior of a phenomenon not in probabilistic notions but in some unnormed weight functions [as Zadeh (1971) does in his theory of fuzzy sets], the statements based on the record, though they will be understood, will have quite another meaning than the statements made in the probabilistic language. Here is an illustration. Assume that somebody, proceeding from certain non-probabilistic considerations, wants to write a formula analogous to the Bayesian one (in the system of notions of subjective probabilities), but for unnormed weight functions

$$p(\mu|y) = kp(\mu)p(y|\mu)$$

If the recorded functions are presented merely as unnormed weight, it will be natural if the coefficient k is to equal 1. Now assume that functions $p(\mu)$ and $p(y|\mu)$ are such that one of them reaches its maximum in one part of the abscissa and the other, in another part, the maximum of one function corresponding to a gently sloping curve of the other one, with ordinate values close to zero. The product of the two functions will yield a bimodal function $p(\mu|y)$ with small weight values for the peaks. We have to acknowledge that the character of the functions $p(\mu)$ and $p(y|\mu)$ makes us give rather an unaccustomed interpretation. It would look as if we are dealing here with a case of "twilight" consciousness when a person cannot clearly enough formulate his ideas. At the same time, if we share probabilistic views, dealing with the same functions $p(\mu)$ and $p(y|\mu)$ we shall obtain a bimodal function $p(\mu|y)$ normalized to 1, which will be familiarly interpreted (in terms of subjective probabilities)

[9] This idea can be illustrated by von Mises's paradox, which was presented earlier (see p. 23). Despite the obvious absurdity of the reasoning in this paradox, the old probability theory lacked anything which would forbid it.

in the following way: in the resulting judgment generated by mixing certain prior information with that received in the given experiment, we have two meanings, and both of them may have approximately equal probabilities. Such conclusions help to explain some seemingly unexpected events of real life.[10]

In criticizing Kolmogorov's axiomatics, Fine pays attention to the fact that two fundamental concepts of probability theory—independence of random values and conditional probability—remain irrelevant to axiomatic constructions: they are given by separate definitions, and Fine believes axiomatics to be incomplete in this respect. But if the structure of probability theory is regarded as the grammar of a language, this remark by Fine, interesting in itself, does not have any essential significance.

Also, if the axiomatics of probability theory is viewed as grammar, the question of its consistency is not of great importance either, and I shall not dwell upon it here[11] [on the growing tolerance to the problem of consistency in mathematics see, for example, Gnedenko (1969)].

In concluding this analysis of the axiomatics of probability theory, I have to acknowledge that not all of its rich content explicated in theorems is used as grammatical structures. Many fairly important theorems of probability theory, e.g., the law of repeated logarithm,[12] have no obvious grammatical interpretation. Mathematical structures, in their practical application, give the language grammar but are not reduced to it.

Physical interpretation of the concept "probability." If the probability theory is considered from a linguistic point of view, then Fine's (1973) complaints, supported by Tutubalin (1972), that from axiomatic struc-

[10] For example, the Bayesian theorem helps to explain the nature of anecdotes in our everyday verbal behavior. Assume that $p(\mu)$ is a prior distribution function of the sense content of a highly polymorphous word. An anecdote may be constructed so that the given word, combined with others, generates in the listener's mind function $p(y|\mu)$ with the maximum in another part of the abscissa, in which the prior distribution function is gently sloping close to the abscissa. As a result, the posterior distribution function will prove bimodal. The anecdotal character of the situation will derive from the fact that the word may have two equally common but essentially different meanings—hence two meanings of the phrase. For more details, see my book *In the Labyrinths of Language: A Mathematician's Journey* (Nalimov, 1981).

[11] It is of interest to observe how the problem is treated in books on probability theory. Tutubalin (1972) formulates it but avoids its detailed discussion, referring only to the fact that the notion of a set, as it is used in constructing axiomatics, leads to paradoxes which cannot be overcome at present in a sufficiently satisfying way. Gnedenko (1969) gives the following argument for the consistency of Kolmogorov's axioms: "Kolmogorov's system of axioms is *consistent* since there exist real objects which these axioms satisfy" (p. 50). Such a basis of consistency, broadly accepted in the pre-Hilbertian period, presupposes acknowledging the above-mentioned postulate by Thomas Aquinas of the World's "consistency."

[12] A remarkable theorem by the Russian mathematician A. Ya. Khinchin specifying the Law of Large Numbers, well known in probability theory. This theorem has brought about a number of serious studies.

ture there does not follow an interpretation of the physical sense of probability, remain incomprehensible.

It is natural to believe—and this is generally accepted at present—that logical grammar deals with symbol systems independently of how they are interpreted in terms of the external word. Interpretation appears later, when language is used to formulate concrete statements. And this interpretation may be polymorphous and fuzzy. Kolmogorov (1956), after his shattering criticism of the conception of von Mises, still gives a frequency definition of probability, though, of course, without transition to the limit. He writes that it suffices to speak of probability as a number around which frequency is grouped under definitely formulated conditions, so that this tendency to grouping is manifested more clearly and accurately with the growing number of tests (up to a reasonable limit).

Such definitions of probability entered the textbooks, too. In Tutubalin's book (1972) we read, "The number around which the frequency of event *A* fluctuates, is called the *probability of event A* and is designated by $P\{A\}$" (p. 6).

We feel a desire to ask: Should this interpretation be considered as the only possible one? It is hard to believe that physicists who study quantum mechanics will agree to this.[13] It is altogether incomprehensible why we should exclude a consideration of probability as a measure of uncertainty in our judgments. If the concept of subjective probability is introduced (as it is by L. J. Savage, Bruno de Finetti, and other representatives of this trend), it proves possible to apply to it all the usual rules of probability calculus.

The requirement of statistical stability. The frequency interpretation of probability immediately gives rise to the problem of stability, very acutely introduced by Richard von Mises. This problem is, if you like, a stumbling-block in discussing all the questions related to the applicability of probabilistic notions for describing external phenomena. On this point, Tutubalin (1972) writes:

> According to modern views, the area of application of probability methods is limited to phenomena characterized by their statistical stability. However, testing statistical stability is difficult and always in-

[13] Here is how the meaning of the wave function is treated by Blokhintsev (1966): ". . . the wave function is not a value determining the statistics of a special measurement; it is a value determining the statistics of a quantum ensemble, i.e., the statistics of any measurements compatible with the nature of microsystem μ and macroscopic situation M which dictates the conditions of movement for the microsystem μ." In his latest book, Blokhintsev (1978) proposes to denote by the term "probability" the measure of the potential possibility of an event's occurring. The American philosopher Abel has collected physicists' statements on the concept "wave function"; in a slightly contracted form it is given in my earlier books (Nalimov, 1974*b*, 1981).

> complete; besides, it often leads to negative conclusions. As a result, in some branches of knowledge, e.g., in geology, it has become a norm not to test statistical stability, which often leads to serious blunders. (p. 144)

In the book by Fine (1973), we find a sad remark that the stability of frequencies, upon which the application of probability theory must be based in the problems of forecasting, in no way follows from Kolmogorov's axiomatics. In some textbooks on probability theory, stability of frequencies is ascribed almost the status of a law of nature. In the book by Ventsel (1962), we read:

> . . . the property of "stability of frequencies," many times tested experimentally and supported by all the experience of human practical activities, is one of the most universal regularities observed in random phenomena. (p. 29)

In the book by Gnedenko (1969), we read,

> Permanent observations over appearance or nonappearance of event A in a large number of repeated tests under the invariable complex of conditions show that for a broad circle of phenomena the number of appearances or nonappearances of event A obeys stable regularities. (p. 41)

I consider all these judgments on the stability of frequency to be a result of misunderstanding, to a certain degree. The concept of frequency stability is nothing more than a logical construction. Without this statement, it is impossible to give a limit-frequency interpretation to the notion of probability. Mathematically, the statement of stability of frequencies is merely a manifestation of the law of large numbers.

This law plays a very important role in the system of probabilistic concepts (for more details, see Gnedenko, 1969). The law allows us to understand (though in a purely logical aspect) why it is possible to use the theory of probability to solve the problems of the real world. But in no way can it serve as a sufficient reason for justifying broad application of theoretico-probabilistic methods since it is very difficult to give a faultless physical interpretation of the conditions which random values must satisfy in order to obey the law of large numbers (for criticism of the law, see Alimov, 1974).

However, nothing definite can be said about the stability of frequencies in the phenomena of the external world, or about statistical stability in a broader sense. There are many real problems in which statistical stability is precisely the object of research, e.g., in the application of analysis of variance in metrological problems to display the dispersion of results of similar measurements taken by various researchers in various laboratories. It is true, however, that the possibility of applying an analysis of variance stems from certain practically untestable prerequisites.

Probabilistic judgments are built, like any other judgment, by proceeding from certain premises. The grammar of probabilistic statements is, generally speaking, nothing more than the rules of constructing grammatically correct phrases (within a given system of concepts) over initial premises. For example, if we study repeated mass phenomena, we can generate grammatically correct (in our system of concepts) judgments concerning the future. But this extrapolation will be legitimate only if the constancy of frequencies is postulated. Information about the constancy of frequences in the future cannot, generally, be obtained from our past experience, nor can we deduce it from the axiomatics of probability theory. Axiomatics only provides us with a grammar that allows us to state what will happen if we accept certain premises.

Now let us consider a slightly different problem. Assume that we wish to predict the future value of the dependent variable from the observational results, using the equation of the straight line. In this case, we obtain the least-squares estimates of the parameters of the straight line; then we build the limits of confidence as two conjugated hyperbolas and make forecasts for the period we are interested in. But, in doing this, we proceed from the following premises:

(1) Errors of estimating the dependent variable are independent random values sampled from the normally distributed universe with a constant, but unknown, variance and with mathematical expectation equaling zero.
(2) Independent variables are estimated without error.
(3) Both parameters of the regression equation have no time drift.

In this case, the forecast is a proposition correctly constructed over these premises, not all of which are of equal importance; some can be slightly violated. Sometimes we even feel in what way the structure of a phrase must be modified if the premises change. For example, if requirement 2 is not fulfilled, regression analysis is replaced by confluence analysis. The most serious requirement is the third, and it is not quite clear whether it can be included among those pertaining to the concept of "statistical stability." One thing is obvious: either the requirement of "statistical stability" should be regarded very broadly, in which case it cannot be introduced into the language's grammar as a separate category, or it may be regarded in a narrow sense, limited, for example, to premise 1, in which case we shall have to stipulate that probabilistic statements should be based not only upon "statistical stability" in a narrow grammatical sense, but also upon stability in a broad sense, which in various problems is displayed in various ways.

In some applied problems, the requirement of "statistical stability" in its narrow sense is not explicitly formulated at all. As an illustration, let us consider a problem of the science-of-science investigated by my post-graduate student S. A. Zaremba. It deals with studying articles cited according to the years of their publication. It has turned out that at the beginning of the eighteenth century, when science, as an information system, was only in the bud, articles cited were evenly distributed according to the years of their publication. To be more correct, it was a mixed distribution composed of several even distributions given at different sections of the time scale and taken with different weights. In the second half of the eighteenth century, the mixed distribution began to contain an exponential constituent. At first it was situated only at the initial section of the time scale and embraced only a small number of publications. But little by little, as we come closer to our time, the greater is the role of the exponential constituent, though the distribution still remains mixed: its tail part preserves the character of an even distribution. The tail part is at present reduced to several dozens of years, whereas at the beginning of the eighteenth century it went as far back as Aristotle's time. The emergence and evolution of the exponential constituent may be interpreted as the representation of the forefront of development in science, which has no roots in the past (i.e., rapidly attenuates in the reverse time direction). Even distribution may be regarded as a particular (degenerate) case of a truncated exponential distribution, which occurred when publications had a relation to all the past experience.

Everything is thus clear. In this research, the concept of distribution functions was used as a specific language cliché to describe a genuinely complicated phenomenon. And we feel that this phenomenon has been aptly clarified by applying familiar stereotypes of the probabilistic language. Nobody worried about the "statistical stability" here. Distribution functions for adjacent time intervals look alike; those for long intervals look essentially different. This was the object of the research. Here, of course, stability was implicit, allowing one to pass from a single observation of frequencies to the concept of probability.

Constructing concepts of probabilistic language. There exist a lot of statements about what mathematical statistics is. I find them interesting to collect and have presented my collection (certainly incomplete) in an appendix to my earlier books (Nalimov, 1974*b*, 1981). In this context it seems pertinent to say that mathematical statistics is a language for constructing statements over values which we like to regard as random.

How was it possible to construct such a language?

Randomness cannot be introduced directly into the system of logical judgments—the latter will immediately prove to be laden with stark contradictions. A system of theoretical constructions had to be formed which would generate concepts enabling the formulation of logically precise descriptions of random phenomena. Among such concepts are general population, sample, probability, distribution function, independent observations, spectral density. These clearly defined ideas and the logical statements built over them are consistent. Randomness has proved to be excluded from the system of logical constructions. It manifests itself only when these constructions are interpreted in the language of experiments, when separate ideas, e.g., mathematical expectation estimated from the sample, are ascribed a fuzzy numerical value, and this fuzzy value is somehow limited by another concept, that of confidence limits. Probability theory, and mathematical statistics in conjunction with it, have reduced the study of randomness to describing random value behavior in probabilistic terms. It has yielded the possibility of describing chance by means of formal logic. The language of such descriptions is weaker than that of causal concepts since it allows us to introduce fuzzy values, at least at the stage of interpretation.

We must be aware of the fact that the concepts of probability theory are certain abstract constructs and not mirror-like reflections of what truly exists in the real world. It is rather a challenge to demonstrate in what way these constructs correspond to what we observe in the real world. There is nothing in the real world to correspond to one of the principal theories of mathematical statistics—that of general population: this concept is a product of profound abstraction. The concept of probability may be shown to correspond to the frequency in the real world if the number of observations is large, though not too large. An overcritical reader will find it hard to understand this. The idea of statistical independence is easily defined in mathematical terms, but it is not so easy to explain to the experimenter how experiments must be carried out in order to get statistically independent results.

On the subject of the difficulty connected with interpreting the term "sample," this is what Tutubalin (1972) says in his brilliant sophism:

> We say that a sample is formed by the results of several independent measurements taken under similar conditions. However, if all experimental conditions are controlled, we shall get one and the same number (there will be no uncertainty), and if not all experimental conditions are controlled, then how do we know that they remain unchanged? (p. 196)

This vagueness of the principal concepts in the sense of their correspondence to reality, of which I could talk much longer, sometimes provocatively gives rise to indignant articles of the kind I have already mentioned (Alimov, 1974). I consider such criticism somewhat illegitimate. One must keep in mind that the language of probabilistic concepts can describe the world only roughly. Let us take the well-known relation

$$\sigma^2\{\bar{y}\} = \frac{\sigma^2\{y\}}{N}$$

In its practical interpretation this is but an approximate relation, and we never know its degree of approximation. The latter is given, on the one hand, by the fact that real observations can never be absolutely independent; on the other hand, it proceeds from the fact that, N being large, experimental conditions no longer remain constant. I believe an experimenter to be statistically educated if he can use the formula wisely. To be able to use mathematical statistics correctly, one has to interpret the limitations formulated in mathematical language in the language of experiment. But, strictly speaking, nobody knows the rules of interpretation.

The requirements which had to be placed upon the behavior of random values while constructing the principal ideas of probability theory proved fairly rigid. It might seem, perhaps, that the real world is more random than is assumed by the language with which we try to describe randomness. Sometimes these requirements may be weakened. Kolmogorov's frequency interpretation of probability mentioned above (see p. 109) is already a weakened (as compared to von Mises's) idea of statistical stability. Indeed, it is impossible to have an infinite series of tests with experimental conditions held constant. Another weakening of requirements made for the behavior of random values is the introduction of *robust* estimates, i.e., those insensitive to initial premises, instead of Fisher's effective estimates, when measuring distribution parameters from samples. As a matter of fact, a grammar of robust estimates cannot be theoretically constructed, so we have to resort to simulating problems in computers to be able to offer recommendations. However, it is also true that the concept of robust estimates cannot be understood without also understanding the concept of effective estimates. For example, the spectral theory of random processes is built only for stationary processes, whereas all, or almost all, observable processes are non-stationary. If the non-stationary aspect cannot be algorithmically removed, then, as follows from the algorithmic probability theory, the non-stationary aspect itself is random. However, nobody can describe this type of randomness; there are no theories within which it could be described. We deal here with a phenomenon generated by a mechanism more complicated than the algorithms we can construct to describe it. In other words, the algorithm for removing the non-stationary aspect cannot be established more compactly than the random sequence itself. One may, certainly, attempt to describe non-stationary processes in the framework of spectral theory, in the manner of Granger and Hatanaka (1964), but such descriptions will seem clumsy.

Every attempt to weaken the requirements imposed upon the behavior of random values by the grammar of statistics irritates mathematicians dealing with probability theory. Laplace, whose contribution to probability theory is remarkable, remained a convinced determinist. Today, too, mathematicians who deal professionally with probability theory and statistics may still share profound and unyielding formalistic views.

In one of the respectable universities of the USSR, the course of mathematical statistics starts off roughly as follows: ". . . 80% of the applications of statistics are wrong since it is applied where there are no random values." In the book by Tutubalin (1972) cited above, we read:

> It is extremely important to eradicate the delusion, sometimes shared by engineers and naturalists insufficiently trained in probability theory, that any experimental result may be regarded as a random value. (p. 166)

So what are the values considered non-random? Those described by causal relations? This is a fallacy. Tutubalin is quite clear on this point, ascribing to non-random values the results of an experiment for which the requirement of statistical stability is not fulfilled. Non-random is what behaves more randomly than is allowed by the language of traditional probabilistic concepts. Is not this notion of randomness obviously inconsistent with its algorithmic definition?

One cannot say how the requirement of statistical stability should be interpreted in each particular case. To be quite meticulous, one will have to limit the applications of mathematical statistics to such experiments as the tossing of a coin and the applications of probability theory to manipulating balls in urns. Even casting dice is not absolutely random because it is not that easy to make perfect dice.

This is precisely what the *art* of statistical analysis consists of: describing in the language of probabilistic concepts the behavior of the real world which is arranged more randomly than is allowed by the grammar of our language. Such a description is sure to be far from successful in many cases.

When the language of probabilistic concepts proves unfit. Sometimes statements formulated in the familiar probabilistic language seem clumsy as a result of the fact that the phenomena described actually reflect truly causal relations that are camouflaged. Here is an illustration from Maslov et al. (1963):

> The problem was to give a statistical foundation for measuring dislocations on the ground edge surface of semi-conductor material. The measurements were taken in the following manner: a net was laid over the ground edge, and the number of dots was calculated which

> were included in the net cells. The results can be well presented as distribution functions. But the latter proved mixed in this case. Their parametric presentation requires computing high order moments, which is clearly inconvenient since it demands a great number of measurements. Besides, the description in terms of distribution functions proves too cumbersome. This is due to the fact that dots on the surface of the ground edge are situated primarily non-randomly, forming clear-cut figures—stars, spears, or merely clouds of condensation. It has been noted that metallurgical engineers can successfully arrange ground edges according to the quality of the material, immediately connecting the etched figures with the physico-chemical properties of the material. The problem thus turned out to be of an obviously topological character and not of a metrical one: it is not the distance among the dots but their entrance into definite sets forming figures which interests the researcher. When this had become clear, it was suggested that the method of evaluating the material's quality be changed. Laboratory assistants were given albums of real and clearly seen etched figures and they were asked to classify ground edges in accordance with the types represented in the albums. This modified method brought about its own statistical problem: it was necessary to estimate how often laboratory assistants at different times classify the same ground edge as belonging to one and the same type.

This phenomenon reflected a causal (though not too prominent) relation between the etched figure and the material's quality, and it was better not to avoid it by a statistical description.

So who can tell in what cases probabilistic language is to be used and when it is not? No general criterion can be proposed. I think it is applicable when the description obtained with its help satisfies us.

Ontology of Chance

What is the physical nature of chance? It seems impossible to answer this question, at least at present.

I shall remind the reader of the way in which the concept of chance is introduced in mathematical literature. In many books on probability theory (e.g., Gnedenko, 1969), the same phrase is repeated, dating back to Aristotle: an event is called random if, under certain conditions, it may or may not happen. The phrase tells us nothing of the physical sense of the concept. The latter is at times linked with generators of randomness, but any such generator produces, among others, sufficiently well-arranged numerical series.

In his well-known book, Hald (1952) makes an attempt to deduce the notion of randomness from that of stochastic independence:

. . . A sample of n observations $x_1, x_2, \ldots x_n$, from a population with distribution function $p\{x\}$ is called *a random sample from that population if*

$$p\{x_1, x_2, \ldots, x_n\} = p\{x_1\}p\{x_2\} \ldots p\{x_n\} \qquad (2.1)$$

It follows that $n!$ different possible orders of given sample values are all equally likely when the sampling is random since the value of $p\{x_1, x_2, \ldots, x_n\}$ is independent of permutations of x's when (2.1) is satisfied. A general *definition* of randomness of a sequence of n observations from the same population in terms of the magnitude and order of these observations therefore seems impossible. (p. 338)

Thus, an attempt to define randomness through stochastic independence proves inconsistent with the notion of randomness which follows from the algorithmic theory where randomness is regarded as a maximal disorder.[14]

The notion of random numbers is, actually, an abstraction. In reality we always deal with pseudorandom numbers, and everybody studying the simulation by Monte-Carlo method knows how cautious one should be concerning the randomness of pseudorandom numbers.

Kolmogorov does not explicitly introduce randomness into his axiomatics. His probability theory is constructed in the framework of a general theory of measure with one special assumption: the measure of the whole space must equal unity.

An algorithmic definition of randomness seems to allow a profound comprehension of randomness from a mathematical standpoint, but it hardly elucidates the physical sense of the concept. It is noteworthy that, being interpreted philosophically, the algorithmic approach to randomness is definition by negation: randomness is defined as something which cannot be described in a deterministic way. It is important to comprehend the significance of this statement thoroughly.

If we turn to philosophical literature, we shall again fail to find fruitful considerations of the ontology of chance. In Soviet philosophy, chance is

[14] It would be of interest here to pay attention to a paradox of randomness in experimental design problems. It would seem natural to consider the experimental design **X** randomly organized if it allows one to obtain stochastically independent estimates of regression coefficients, i.e., such estimates for which cov $\{b_i b_j\} = 0$. In this case all the non-diagonal elements of the information matrix $\mathbf{X}^x\mathbf{X}$ should equal zero. But such a design can be built by using, say, a Hadamard matrix. This is a square matrix of the order N consisting of the elements $+1$ and -1 and having the property that $\mathbf{X}^x\mathbf{X} = N\mathbf{I}$. From the definition it follows that all the non-diagonal elements equal zero and, consequently, all the covariances for the regression coefficients estimates also equal zero. If now we try to construct an experimental design of the same dimension randomly placing $+1$ and -1 in the cells of the table, then we, as a rule, obtain designs for which non-diagonal elements will be comparatively small but not equal to zero. It turns out that, at least in some cases, regular modes of construction yield an experimental design generating regression coefficients estimates arranged "more randomly" than designs built randomly. Recently, it became known that to construct a valid sequence of random numbers (satisfying many criteria) one should use not random procedures but some rational ones.

now elevated to the rank of a philosophical category, which, of course, is an important rehabilitation of the concept. But in reality this leads to the following:

> Randomness—a kind of relation determined by external causes secondary for a given phenomenon or a process. Random relations are characterized by unstable and temporal occurrences, relative indifference towards the form of its manifestation, and uncertainty of emergence in space. The category of randomness is correlated to that of necessity. (Yakhot, 1970, p. 33)

I am sure physicists will reject this definition of randomness. Are we to say that the movement of gas particles, experimental errors of measurements, radioactive decay, and probability waves in the microworld are determined by secondary causes? If these secondary causes are removed, experiments will become free from error, radioactive decay will lose its random character, and the concept of probability waves will merely disappear. Mathematicians dealing with probability theory will be even more indignant: they require statistical stability, and the definition says that random relations are characterized by their unstable character.

But let us leave philosophers alone and turn our attention to the literature of popular science. This kind of literature is interesting in that it reflects only what is indubitably acknowledged by the existing paradigm. On the table in front of me there are several books of this kind that have a direct bearing on the matter in question. One of them is the book *This Random, Random, Random World* by Rastrigin (1969). In it we read:

> Indeed, any event has a quite definite cause, that is, is an effect of this cause. Any random event has such a cause, too. (p. 5)
>
> . . . randomness is, first of all, . . . unpredictability resulting from our ignorance, our insufficient knowledge, lack of necessary information. (p. 8)

This sounds like an age-old incantation: "I am not a heretic; I do believe in causality!" However, later on the author has to make concessions and to speak of the uncertainty principle, inexhaustability of the universe and limited human possibilities—in brief, to state the impossibility of getting rid of randomness. So what is randomness, then? Is it only our ignorance and something unique in the microworld?

In the book *Natural Philosophy of Cause and Chance* by Born (1949) we read:

> The notions of cause and chance which I propose to deal with . . . are not specifically physical concepts but have a much wider meaning and application. . . . It would be far beyond my abilities to give an account of all these usages, or to attempt an analysis of the exact significance of the words "cause" and "chance" in each of them. . . .

> Indeed, cause expresses the idea of necessity in the relation of events, while chance means just the opposite, complete randomness. Nature, as well as human affairs, seems to be subject to both necessity and accident. Yet even accident is not completely arbitrary, for there are laws of chance, formulated in the mathematical theory of probability. . . . In fact, if you look through the literature on this problem you will find no satisfactory solution, no general agreement. Only in physics has a systematic attempt been made to use the notions of cause and chance in a way free from contradictions. (p. 1)

Further, Born prefers to speak of the concrete sense the concept of chance has in various physical problems.

We cannot learn what randomness is from the book *Causality and Chance in Modern Physics* by Bohm (1957) either. We read there:

> Indeed, the laws of chance are just as necessary as the causal laws themselves.[1] For example, the random character of chance fluctuations is, in a wide variety of situations, made inevitable by the extremely complex and manifold character of the external contingencies on which the fluctuations depend. . . . Moreover, this random character of the fluctuations is quite often an inherent and indispensable part of the normal functioning of many kinds of things, and of their modes of being. (p. 23)
>
> [1] Thus, necessity is not to be identified with causality, but is instead a wider category.

Thus, we learn that randomness is inherent to nature and is part of necessity. Probably all this saves us from heresy, making us believe that everything is necessary, but this is hardly essentially elucidating.

In the book by Blokhintsev (1966), His Majesty Chance is introduced without any incantations, and his role in the quantum-mechanical conception of the microworld is described. I feel that nothing better can be done.

Any attempts to comprehend the ontology of chance lead to obviously superfluous statements. It seems better to say that randomness is not an ontological cateogry, but an epistemological one. Or that, like necessity, this is one of the two categories generating two languages for describing the world. In both cases, we deal not with concepts emerging as a mirror-like reflection of reality but with abstractions built over the observed external world, abstractions generating two different grammars for arranging and comprehending our observational results. Here we cannot but recollect Bohr's principle of complementarity. If this linguistic viewpoint is acknowledged, we immediately succeed in climbing out of the bog of reasonings on chance's ontology and become free from the need to sing incantations. The highly readable collection of papers *Sovremennyi Determinizm* (Svechnikov, 1975) is a fine example of the difficulties one faces in attempting to ascribe an ontological sense to the ideas of causality and chance.

To my mind, the failure of all the attempts to comprehend chance ontologically has a very simple explanation: their aim is to achieve the impossible, i.e., to explain chance in the familiar framework of deterministic ideas.

The idea of the ontological meaning of the concept of chance can be successfully grounded only within the extreme philosophical manifestations of ideas which are customarily referred to as irrationalism. It is noteworthy that the acknowledgment of ontological chance is accompanied by the rejection of ontological causality. I shall illustrate this by a quotation from Sartre's (1965) famous *Nausea*:

> The essential thing is contingency. I mean that, by definition, existence is not necessity. To exist is simply *to be there*; what exists appears, lets itself be *encountered*, but you can never *deduce* it. There are people, I believe, who have understood that. Only they have tried to overcome this contingency by inventing a necessary, causal being. But no necessary being can explain existence: contingency is not an illusion, an appearance which can be dissipated; it is absolute, and consequently perfect gratuitousness. (p. 188)

This elegant statement looks very pertinent in the system of Sartre's existentialism. But here we try to remain within the frame of scientific reasoning.

The description of phenomena in terms of chance makes the world more mysterious than the determinist believes it to be. As a matter of fact, this is a purely psychological effect which disappears during a subsequent logical analysis. Indeed, consistent determinism makes us acknowledge certain initial causes, such as laws of nature which had emerged without cause.[15] Mysteriousness, inherent in determinism, is merely shifted to the remote past. In the system of deterministic ideas, the world, emerging without cause, now proves causally arranged, and probabilistic notions destroy the arrangement and introduce the absence of cause into the description of our everyday experience.

Finally, I would like to show how the phenomena which cannot be described in the framework of causal concepts can be described with the help of the concept of chance. Imagine a physical apparatus registering

[15] The statement that initial causes never appeared but had always existed is nothing more than the acknowledgment of difficulties which arise while unwinding the chain of causal links, in a program extrapolating into the past. Actually, it is hard to imagine how something absolutely unchangeable and uncreated which, nevertheless, generates a changing world can exist in time. All this resembles theological structures in which consistent determinism unavoidably leads to the perennial First Cause. But within well-reasoned philosophical-religious systems, e.g., in gnosticism, it was at least stated that the initial cause, God, exists outside time. Moreover, God was sometimes described as "non-existing"; otherwise, one had to look for the cause of his appearance.

the radioactive decay of atoms at some moment of time. What determines the process for a given atom at a given moment of time? Modern physics does not answer this question: one has to acknowledge that the search for the so-called "latent parameters" is clearly a hopeless task (see, for example, Svechnikov, 1975). It may be said that the decay of radioactive atoms obeys statistical regularities. But the sense of this statement lies in the fact that the actually observed frequencies, for some reason or other, behave in such a stable manner that it becomes possible to speak of presenting observational results by distribution functions. The knowledge of their parameters allows us not only to forecast the process of decay in time with great certainty, but also to control many physical experiments. Here we deal with a description which makes it possible to master nature without penetrating into the essence of the phenomena. One can certainly say that statistical regularities are a special case of a broadly understood principle of determinism. Hence, it would seem to follow that, by force of necessity, the given atom does undergo decay at a certain fixed moment of time. However, this statement will hardly differ from the statement that the decay of the atom was caused by the will of the Demiurge, the creator of worlds. In neither case can we support our statements by any substantial data. We cannot acknowledge, even in a purely speculative way, that this decay fixed in time was caused by a trigger, something like an alarm clock randomly set in the infinite past.

Any reasoning of this sort makes things more puzzling rather than clarifying them. Still, this is not to say that we are going to share the viewpoint of agnosticism. We are only made to acknowledge that we have to use a language containing concepts whose physical sense, after serious consideration, proves to be fairly vague. The odd fact is that it is with the help of such concepts that the world is described and mastered. This is an amazing peculiarity of our scientific language. Why should we not discuss these matters directly?

And now a few words about the well-known book *Chance and Necessity* by Monod (1972). In discussing Darwin's evolutionary theory, Monod remarks that the prevailing importance should be attached to the variability at the molecular level rather than to the struggle for existence, an idea that belongs not to Darwin but to Herbert Spencer. However, variability related to the molecular interpretation of Darwinism can be described only in terms of chance: its inexhaustible resources have to be connected, according to Monod, with the ocean of chance. This, in its turn, results in difficulties of comprehension:

> Even today a good many distinguished minds seem unable to accept or even to understand that from a source of noise natural selection alone and unaided could have drawn all the music of the biosphere. (Monod, 1972)

Monod says further that the progress of evolution comes from external conditions that place limitations on chance. But the basic thing is still variability generated by chance. He remarks that not only phylogenesis has to be described in terms of chance, but some local phenomena as well—e.g., the process of formation of specific antibodies for destroying newly emerged antigens. No information is borrowed from antigens when antibodies are synthesized. Everything happens as in playing roulette.

My first reaction to Monod's book was a feeling of sadness. We are so accustomed to perceiving the world in terms of causal concepts that the description in terms of chance seems to lack an explanatory power. However, by and by, a new sensation arose—the impression that biology is now facing a revolution probably even more crucial than that in twentieth century physics. We come to understand that the mystery of life, as well as the mystery of the microworld, can be described only in terms quite new and unfamiliar. The world confronted by modern science proves so complicated that it cannot be described in the familiar system of ideas. To describe this complexity, we had to invent a new language containing concepts with a vague physical sense. I should probably add that the physical meaning of these concepts is unclear because of our desire to comprehend them within the system of old ideas.

Concluding Remarks

Let us try to sum up. Determinism is deeply rooted in the history and pre-history of human thinking. The concept of chance evidently appeared much later, when it was understood that the search for the causal explanation of all phenomena inevitably leads to fantastic conceptions. However, it took a long time to coordinate randomness with a formal, logical way of constructing judgments, and European philosophical thought, both scientific and religious (they were quite in agreement on this question), spent ages to trying to avoid randomness by explaining it simply as insufficient knowledge. Probability theory, having laid serious limitations upon manifestations of randomness, created a language that allowed us to describe the latter within strictly logical structures. This language proved richer than that of rigid determinism and gave the opportunity to describe phenomena in a fuzzy manner without arranging them in a system of rigid causal relations. The position of determinism grew weaker. The most extreme position among probabilistic trends is occupied by the school of subjective probabilities, based upon the neobayesian approach. The prior distribution function can be regarded here as a system of fuzzy (probabilistically weighted) axioms, and the posterior distribution, as a fuzzy judgment. Not only among philosophers, but also among mathematicians who study probability theory, the struggle continues against the acknowledgment of chance: some of the latter are trying to limit probabilistic concepts to an extreme formalism.

Philosophically disposed thinkers like to ask whether there is progress in the history of human thinking. Scientific achievements can hardly be claimed to be the best manifestation of progress. Pragmatically, they certainly have made life easier, but they also have brought mankind face to face with the threat of ecological catastrophe; epistemologically, all scientific results can be interpreted as nothing more than mastering nature, since our knowledge of today, from the point of view of tomorrow, is only paradigmatically fixed ignorance. It is not the change of ideas that matters, but the evolution of human thinking. To this extent, we *have* progressed in our comprehension. One can actually speak only of this progress in thinking. We have to acknowledge that, as science develops, thinking does grow broader. The constricting framework of dulling determinism is collapsing, though from time to time we observe efforts to save it, by implicit or explicit recognition of logical positivism. The acknowledgment of chance is not the only attempt directed at broadening our thinking. Other attempts to manifest the freedom of thinking can be indicated. These include Bohr's principle of complementarity, to which its author endeavored to ascribe a universal character, and attempts to construct a many-valued logic, in particular, the three-valued logic of Reichenbach, designed to formalize physical theories. However, not all of these attempts are welcomed by everyone.

I do not ask the reader to turn to irrationalism. The problem can be solved, at least partially, by weakening formal logic. Without logic we cannot say anything coherent. In my earlier book *In the Labyrinths of Language: A Mathematician's Journey* (Nalimov, 1981), I tried to use the neobayesian approach to explain the irregularity of our verbal behavior, and attentive readers could not but notice that my reasoning was based on common logic. The same rebuke was made by Born to Reichenbach when the latter developed his concept of three-valued logic. Born (1949) wrote:

> Concerning the logical problem itself, I had the impression while reading Reichenbach's book that in explaining three-valued logic he constantly used ordinary logic. This may be avoidable or justifiable. (p. 108).

The same is true of everything written above. It's up to the reader to judge!

Back to Contents
Next to Chapter 5