ELSEVIER

Theoretical Computer Science 207 (1998) 263–317

Theoretical Computer Science

Mathematical metaphysics of randomness[1,2]

Andrei A. Muchnik[a,*,3], Alexei L. Semenov[a], Vladimir A. Uspensky[b,3]

[a] *Institute of New Technologies, 10 Nizhnaya Radishevskaya, Moscow 109004, Russia*

[b] *Department of Mathematical Logic and the Theory of Algorithms, Faculty of Mechanics and Mathematics, Lomonosov Moscow State University, Vorobjyovih Gorih, V-234 Moscow, GSP-3, 119899 Russia*

Abstract

We consider various mathematical refinements of the notion of randomness of an infinite sequence and relations between them. On the base of the game approach a new possible refinement of the notion of randomness is introduced – unpredictability. The case of finite sequences is also considered. © 1998—Elsevier Science B.V. All rights reserved

Keywords: Random sequences; Lawless sequence; Stochastic sequence; Unpredictable sequence; Chaotic sequence

Contents

* Corresponding author. Tel.: 7095 9151394; fax: 7095 9156963; e-mail: amuchnik@int.glasnet.ru.

[1] The present authors gave a joint talk entitled Randomness and Lawlessness at the IMSSS & CSLI Conference on the Foundations of Randomness which proceeded from 4 to 7 March 1991 at the Stanford University, Stanford, CA. This paper is an extended version of that talk.

[2] The work was supported in part by the Institute of New Technologies, Moscow, Russia.

[3] Work supported in part by the INTAS project No 93-0893.

PII S0304-3975(98)00069-3

0. Introduction

The mathematical formalization of probability theory was completed in the 1930s by Kolmogorov. Let us mention that it was hundred years later than Cauchy's papers had given the mathematical formalization for mathematical analysis. The century delay is due to the complexity of the notion of probability and related notions. Indeed, probability laws are concerned with the future and not with the past, i.e., with events that may happen in the future rather than with events that have happened. The distinction between the past (what has already happened), the present (the moment of observation and prediction), and the future (what may happen) is crucial in the ideological basis of probability theory. The opposition between *constatation* of what has happened and *prediction* of what may happen is the special feature of probability theory, and it is present in no other branch of mathematics in such a strong form.

Philosophical arguments on the nature of probability last till today. It is not easy, thinking on probability, to distinguish between objective and subjective. Von Mises pointed out this difficulty, when he stated that the absence of knowledge about the true state of affairs cannot be a base of the hypothesis of equal chances. Von Mises wrote [15] (the citation is from the English translation [16, p. 75]):

"If we know nothing about the stature of six men, we may presume that they are all of equal height. $\langle\cdots\rangle$ This presumption may be true or false; it can also be described as more or less probable, in the colloquial meaning of this word. In the same way we can presume that the six sides of a die, of whose properties we know nothing definite, have equal probabilities. This is, however, only a conjecture, and nothing more. Experiment may show that it is false, and the pair of dice used in our first lecture was an illustration of such a case".

Paradoxes in probability theory are well known. A common example is: we easily agree with the following outcome of 20 trials of fair coin tossing: 01111011001101110001, on the other hand, we wonder the result 00000000000000000000. However, both results have equal chances. It is plausible that our reaction is caused by our belief that the first string is complicated, i.e., has only complicated descriptions and the second string is simple, i.e. has a simple description (the question if, in a random experiment, a string which has small probability and is simple may occur at all, needs a further discussion). The paper [22] starts with an example of that kind; the problem is discussed in the special Addendum in that paper, called "A timid criticism regarding probability theory". We will return to this discussion in the end of the present paper (see Philosophical Supplement).

This paper consists of 12 sections.

In Section 1, we fix certain general ideas on randomness regarding infinite binary sequences to try, in the next sections, to find mathematical formalizations of those ideas, based on the theory of algorithms. The first attempt is made in Section 2; there, the rigorous notion of a lawless sequence is proposed as such a formalization; however, we discover fast that formalization is not appropriate. In Section 3, we consider the frequency approach to randomness, going back to R. von Mises; the resulting rigorous notion of a random sequence is called *stochasticness*. In Sections 4 and 5, certain principles are stated, which, we believe, any appropriate formalization must satisfy.

In Sections 6 and 7, an approach to randomness based on the notions of a game and of a strategy is investigated; we declare that random sequences are exactly those against which we cannot win in the game. As a mathematically rigorous notion reflecting that approach, the notion of an unpredictable sequence is defined. That notion turns out to be more restrictive than the notion of stochasticness (7.4 and 7.7).

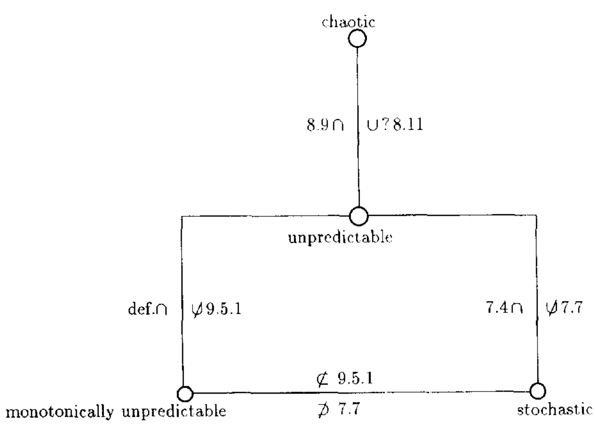
In Section 8, the complexity approach (more exactly, the entropy approach) to the notion of a random sequence is discussed. That approach is due to A.N. Kolmogorov. There, *chaotic* sequences are considered, that is, those sequences whose initial segments' entropies grow sufficiently fast. All chaotic sequences turn out to be unpredictable (8.9), the truth of the converse is an open question.

In Section 9 we find out how fast the entropies of initial segments of a sequence which has certain features of randomness may or must grow. There, the role of *monotone* strategies is clarified and the important notion of a *natural* sequence is introduced.

In Section 10, the ideas of the previous sections are applied to tuples, or finite sequences. From the point of view of practice, as well as of philosophy, it is the investigation of randomness of finite sequences that has the main interest. Infinite sequences have to be considered as upper approximations to finite sequences. However, dealing with finite sequences meets considerable difficulties. As to the class of infinite sequences, we may hope to divide it into two parts – the subclass of random sequences and the subclass of non-random sequences. Such a partition is impossible for the class of finite sequences; thus in this case we have no choice but to evaluate the amount (or the degree) of randomness of a finite sequence.

The questions which remain still open are listed in Section 11 and Section 12 is a philosophical supplement.

For convenience of the reader we draw a pictorial representation of relations between classes of infinite sequences corresponding to four main refinements of the notion of randomness: chaoticness, unpredictability, monotonic unpredictability and stochasticness.



We write the numbers of propositions where the corresponding inclusions and non-inclusions are formulated. The question, whether the inclusion of the set of chaotic sequences in the set of unpredictable sequences is strict, is open.

A brief summary of the notations followed in this paper is listed below.

$\mathbb{N}$	the set of natural numbers (i.e. of positive integers)
$\mathbb{Q}$	the set of rational numbers
$\mathbb{R}$	the set of real numbers
Ω	the set of infinite binary sequences
Ξ	the set of binary tuples, or of finite binary sequences
Ω_s	the set of infinite binary sequences that extend the tuple s
I^m	the set of binary tuples of length m
$I^{\leqslant m}$	the set of binary tuples of length not greater than m
I_s^m	the set of binary tuples of length m that extend the tuple s

If $s=\langle a_1,\dots,a_n\rangle$, $b\in\{0,1\}$ then $s\hat{\ }b$ denotes the tuple $\langle a_1,\dots,a_n,b\rangle$

If $s=\langle a_1,\dots,a_n\rangle$, $t=\langle b_1,\dots,b_k\rangle$ then $s\hat{\ }t$ denotes the tuple $\langle a_1,\dots,a_n,b_1,\dots,b_k\rangle$

If $\mathbf{a}=\langle a_1,a_2,\dots\rangle$ is a tuple or an infinite sequence then $\mathbf{a}|n=\langle a_1,a_2,\dots,a_n\rangle$, and $\mathbf{a}(j)=a_j$

$\mathrm{lb}(x)$	the logarithm of x to the base 2
$\ln(x)$	the logarithm of x to the base e
$\lvert A\rvert$	the cardinality of A (if A is a set)
$\lvert s\rvert$	the length of s (if $s\in\Xi$)
$\mathrm{lh}(s)$, $\mathrm{l}(s)$	the length of s
η	the uniform Bernoulli distribution, or the uniform Bernoulli measure
$[x]$	the greatest integer less than or equal to x

1. The intuitive concept of randomness

The concept of randomness arises from the viewpoint on random things as on things that have unpredictable behavior. The most fruitful way of converting intuitive concepts of such a kind into exact mathematical notions is their application to a class of phenomena having the rigorous mathematical description. Coin tossing produces such a class of phenomena, so we will consider the concept of randomness for infinite binary sequences.

The restriction to binary case is certainly strong, however all facts discovered for that case can be extended via coding to the sequences of objects of an arbitrary nature (though some quantitative laws may be slightly changed).

Before further exposition we would stress on the following point. Our opinion about the randomness of a sequence depends on accepted probabilistic model, which must be chosen in advance. The following simple example illustrates this point. Let us toss a symmetric coin and let us consider a sequence of outcomes of mutually independent trials of the experiment. We expect that in every long initial segment of the sequence there are ones about as many as zeros, and if the number of zeros is sufficiently greater than the number of ones we consider this fact as an evidence of nonrandomness of the sequence. If our coin is not symmetric (for example, one side has probability $\frac{2}{3}$ and the other side has $\frac{1}{3}$) and the sequence has ones about as many as zeros then we consider it as non-random: in any random sequence there are zeros about twice as many as ones.

In further exposition we will consider symmetric coins as well as asymmetric coins, for which the probability p of one is not equal to the probability q of zero (in some cases these probabilities will depend on the number of a trial; more complicated probability distributions will also be considered). However we will point out every case of using an asymmetric coin or a coin with changing probabilities or even with a more complicated distribution. If it not stated otherwise all the trials are mutually independent and the probabilities of one in each trial is equal to $\frac{1}{2}$.

Remark. It is known that in any mathematical treatment of randomness, an infinite sequence having positive measure is treated as random with respect to the measure in consideration. (Speaking on the measure of a sequence we have in mind the measure of the single-element set consisting of that sequence.) Therefore every sequence may be random with respect to an appropriate non-continuous measure. This remark can be applied also to any refinement of the notion of randomness being developed in this paper – to typicalness, to chaoticness, to unpredictability, to stochasticness. So the reader must not be confused when he realizes that, say, the sequence $\langle 0, 0, 0, \ldots \rangle$ containing zeros only is unpredictable with respect to some exotic measure.

2. A first attempt to develop the concept of a random infinite sequence: lawfulness and lawlessness

We would define randomness as the absence of any law. Thus we need to make clear what a law is. Generally, this means that the terms of a given (binary) sequence cannot be mutually independent, or more exactly, lawfulness is the presence of such dependence. The most easy kind of such dependence is the existence of some nontrivial information about a term of the sequence in its preceding terms.

We can accept that this information is nontrivial in every step i.e. for every term; this assumption is our first approach.

Definition 2.1. An infinite binary sequence is *lawful* if there is an algorithm that given any initial segment of the sequence outputs a nontrivial finite set of finite extensions of that segment such that one of these extensions is an initial segment of the sequence. Any infinite binary sequence which is not lawful is, by definition, *lawless*.

One may believe that as lawlessness reflects some features of randomness, so our definition of a lawless sequence may serve as an approach to the concept of a random sequence. We shall see, in the end of this section, that this is not true.

Now we will give some terminological remarks and comments. Finite sequences are also called *tuples*. An *initial segment* of infinite sequence $\mathbf{a} = \langle a_k : k \in \mathbb{N} \rangle$ ($\mathbb{N}$ stands for the set of natural numbers, which begins with 1) or of finite sequence $\mathbf{a} = \langle a_k : 1 \leqslant k \leqslant s \rangle$ is, by definition, any tuple of the form $\mathbf{a}|n = \langle a_1, \ldots, a_n \rangle$, that is any tuple consisting of initial terms of the sequence (here $n \leqslant s$). In this case we say that the finite or infinite sequence $\mathbf{a}$ is an *extension* of the tuple $\mathbf{a}|n$, or that $\mathbf{a}$ *extends* $\mathbf{a}|n$. If $s = \langle s_1, \ldots, s_n \rangle$, $s' = \langle s_1, \ldots, s_n, s_{n+1} \rangle$ (thus $s = s'|n$) then we write $s' = s\hat{\ } s_{n+1}$.

The number of terms of tuple b is called its *length* and is denoted by $\mathrm{lh}(b)$ or $\mathrm{l}(b)$ or $|b|$. In particular, $\mathrm{lh}(\Lambda) = \mathrm{l}(\Lambda) = |\Lambda| = 0$, where Λ stands for the empty tuple.

In the sequel the infinite sequences will be called simply *sequences*.

We will consider *binary* sequences and tuples, that is sequences and tuples consisting only of zeros and ones. The set of all binary tuples is denoted by Ξ, the set of all infinite binary sequences is denoted by Ω. We will denote by Ω_b the set of all infinite binary extensions of tuple b. Any set of the form Ω_b is called a *cone*, or an *interval*, or a *sphere* in Ω. Evidently a tuple b is an initial segment of a tuple c if and only if $\Omega_c \subseteq \Omega_b$.

Obviously, there is a computable bijection from $\mathbb{N}$ onto Ξ, thus the notion of partial recursive function can be naturally extended on partial mappings from Ξ into $\mathbb{N}$, $\mathbb{N}$ into Ξ and Ξ into Ξ and on the partial mappings from Ξ into the set $\mathbb{F}$ of all finite subsets of Ξ (obviously, there is a computable bijection from $\mathbb{N}$ onto $\mathbb{F}$). In our exposition we make no difference between the notions “algorithm”, “computable function”, “partial recursive function”.

The algorithm from the Definition 2.1 given any initial segment $\mathbf{a}|n$ outputs a finite non-empty set $f(\mathbf{a}|n) \subseteq \Xi$ such that every tuple in $f(\mathbf{a}|n)$ extends $\mathbf{a}|n$. The nontriviality required in the definition means that the tuples from $f(\mathbf{a}|n)$ do not cover $\mathbf{a}|n$; more exactly, the intervals Ω_t, $t \in f(\mathbf{a}|n)$, do not cover the interval $\Omega_{\mathbf{a}|n}$, that is there is a sequence extending $\mathbf{a}|n$ but extending no tuple from $f(\mathbf{a}|n)$.

We say that a sequence $\mathbf{a}$ is *lawful via* an algorithm f if this algorithm f is defined (that is it outputs a result) at least on all tuples of the form $\mathbf{a}|n$ and for any n:

(1) $f(\mathbf{a}|n)$ is a nontrivial finite set of extensions of tuple $\mathbf{a}|n$, and

(2) there is a tuple $t \in f(\mathbf{a}|n)$ which is an initial segment of the sequence $\mathbf{a}$.

Now it is clear that a sequence is *lawful* if and only if it is lawful via some algorithm.

To make more clear the nature of lawfulness it is convenient to consider the following property of a sequence that is equivalent to lawfulness. We need canonical topology on the set Ω of all infinite binary sequences (Cantor’s discontinuum) the base of that topology consists of all sets of the form Ω_t, i.e. of all intervals. A set open in this topology is called *effectively open* if it can be represented as a union of recursively enumerable set of intervals.

Theorem 2.2. *A sequence* $\mathbf{a}$ *is lawful if and only if it belongs to the boundary of some effectively open set.*

Proof. Let a sequence $\mathbf{a}$ be lawful via an algorithm f. Let us construct a new (possibly not totally defined as well as f is not) algorithm $g : \Xi \to \Xi$. Let $s \in \Xi$. If $f(s)$ is defined and is a nontrivial set of finite extensions of s then there is a tuple r extending s such that the length of r is greater than the length of any tuple in $f(s)$ and r extends no tuple in $f(s)$; the first such r (we assume that some effective enumeration of Ξ is fixed) is taken as $g(s)$. If the condition of the above sentence is not fulfilled then $g(s)$ is not defined. The set $G = \bigcup_s \Omega_{g(s)}$ (the union is over all s such that $g(s)$ is defined) is effectively open. Let us verify that **a** belongs to the boundary of G. This means that in any neighborhood of **a** there are points from G as well as points that do not belong to G. By the choice of f and by the construction of g and G we have $\mathbf{a} \notin G$. On the other hand, every neighborhood $\Omega_s = \Omega_{\mathbf{a}|n}$ of the point **a** intersects with G because $\Omega_{g(s)} \subseteq \Omega_s \cap G$.

Conversely, let a set $G \subseteq \Omega$ be effectively open and let **a** belong to the boundary of G. Obviously $\mathbf{a} \notin G$ as G is open. Let $G = \cup\,\Omega_t$, $t \in M$, where M is a recursively enumerable set. For every s let us denote by $g(s)$ the first tuple $t \in M$ (in the recursive enumeration of M) extending s (if there is no such t then $g(s)$ is not defined). Let us define $f(s)$ to be the set of all extensions of s having the same length as $g(s)$ has except of $g(s)$. One can easily verify that **a** is lawful via f (this is left to the reader). □

Let us recall that a *meagre* set (or a set of the *first category*) is any countable union of nowhere dense sets. Recall that a set $X \subset \Omega$ is *nowhere dense* if every interval includes another interval that does not intersect with X, i.e. if every tuple s has an extension t such that $\Omega_t \cap X$ is empty. We can easily verify that the boundary X of any open set $Y \subseteq \Omega$ is nowhere dense (if $\Omega_s \cap Y = \emptyset$ then Ω_s has no points from X, thus we can take s as t; and if Ω_s intersects with Y then $\Omega_s \cap Y$ includes an interval Ω_t which evidently has no points from the boundary).

Thus we have

Corollary 2.3. *The set of all lawful sequences is a set of the first category.*

Our definition of lawfulness and lawlessness is based on the idea which is well known in mathematics. For example, *generic* sequences defined in the branch of axiomatic set theory called forcing method (see [1, ch. IV]) are lawless in a sense close to our sense.

However, the concept of lawlessness cannot be considered as satisfactory approach to the concept of randomness because the coin tossing yields *lawful* sequence which we must consider as random. Indeed, any infinite sequence of outcomes of mutually independent trials of tossing a symmetric coin satisfies the law of large numbers (LLN): the ratio n_1/n, where n_1 is the number of ones in the sequence's initial segment of length n, tends to $\frac{1}{2}$ as n tends to infinity. But

every sequence **a** *satisfying* LLN *is lawful.*

Indeed, let a natural number N be so large that $n_1/n > 0.4$ for all $n > N$. Let us define an algorithm f such that **a** is lawful via f. Let $s \in \Xi$ be a tuple of a length $m \geqslant N$. Let us denote by $f(s)$ the set of all tuples t of length $3m$ extending s such that the number of ones among the members of t is at least $0.4 \times 3m$. Let us additionally define

$f(\mathbf{a}|k)=\mathbf{a}|(k+1)$ for $k<N$ (thus f is defined on all initial segments of $\mathbf{a}$). Evidently $\mathbf{a}$ is lawful via f.

Analogously, we can verify that any sequence of outcomes of tossing an asymmetrical coin, with different probabilities of one and zero, is lawful.

More generally, sequences obtained by coin tossing, which we consider as random, are lawful because they *satisfy the laws of probability theory* that seems to be necessary for random sequences. Therefore the notion of lawlessness cannot be an adequate approach to the concept of randomness, and we have to find another way to clarify that concept, a way which would be more adequate to our intuition.

3. The frequency approach to the concept of randomness: stochasticness

Thus the law of large numbers makes it impossible to interpret randomness in pure negative sense, i.e. as lawlessness. Let us look whether it is possible to obtain a positive effect from this observation: if the (intuitively) random sequences satisfy LLN (and probably another laws) then it seems quite natural to declare this law (probably together with another laws) to be the criterion of randomness.

However, LLN for the whole sequence only is not enough because the sequence $\langle 0,1,0,1,0,1,\dots\rangle$ certainly is not random and satisfies LLN. Taking this into account, von Mises [19, 15] suggested to require that not only the sequence but also its subsequences must satisfy LLN. Of course we cannot understand this literally (let us take the subsequence of zero terms); however we can extract from this proposal a correct idea. Namely, we must restrict the ways of choosing the subsequence, in particular, by requiring the way of choosing to be effective (that is the way must be represented by an algorithm). On the other hand, it is appropriate to consider first of all only the ways of choice of a subsequence according the following framework: one chooses the terms in turn and the decision of including the current term into the subsequence does not depend on the value of current term (i.e. the term must be chosen before one gets its value).

For example, the following way of choosing does not satisfy the latter requirement: to choose all terms such that the preceding and succeeding terms are equal to zero (i.e. a_n is chosen in the subsequence iff $a_{n-1}=a_{n+1}=0$): there is no way to know whether a_n will be included in the subsequence before getting the value a_{n+1}. It is clear that the subsequence of the sequence $0,1,0,1,0,1\dots$ (evidently satisfying LLN) chosen according the above rule will not satisfy LLN. (However, we can prove that for almost all sequences $\mathbf{a}$ with respect to the uniform measure η the subsequence chosen according to this rule satisfies LLN.)

The most general framework of formalization of the frequency approach was given by Kolmogorov [5, Remark 2]. To this end he gave the notion of *admissible rule* of choice. Let us give an account of his definition. Let us imagine that Man step by step and according to some *rule of choice* points out the number of a term the value of which he wants to know. Let us require that before its (correct) value becomes known to Man[4] he must decide whether he will include this term in the subsequence. All information that Man possesses is the sequence of values of known terms.

The rule of choice (or the *place selection rule*) may be not totally defined, i.e. in some cases it may give no recommendation regarding what should be done – in this case we get a finite subsequence.

Warning. We use the term "subsequence" in a non-standard sense. Usually the term "a subsequence of a sequence $\mathbf{x} = \langle x_1, x_2, \ldots, x_m, \ldots \rangle$" means any sequence of the form $\langle x_{m_1}, x_{m_2}, \ldots, x_{m_k}, \ldots \rangle$ where

$$m_1 < m_2 < \cdots < m_k < m_{k+1} \cdots \qquad (*)$$

Following Kolmogorov, we reject the condition $(*)$ and accept that the terms of a subsequence can be arranged in any way. For example, if we rearrange in an arbitrary way the terms of the sequence $\mathbf{x}$ then the obtained sequence is considered to be a subsequence of $\mathbf{x}$. By definition, if we apply a rule of choice to a sequence then the chosen terms are enumerated, in the capacity of terms of the subsequence obtained, according to the order in which they are chosen (not according to the order in which they are arranged in the given sequence!).

Definitions 3.1. Let Θ be a rule of choice and $\mathbf{a}$ be a binary sequence. Assume that after application of Θ to $\mathbf{a}$ we get a finite sequence or we get an infinite sequence satisfying the law of large numbers, i.e. the limit of the ratio (number of ones in the initial segment)/(the length of the segment) equals to $\frac{1}{2}$ (if the obtained sequence is finite we demand nothing). In this case we will say that $\mathbf{a}$ *is stochastic with respect to* Θ. If the obtained sequence is infinite and does not satisfy the law of large numbers then we say that $\mathbf{a}$ is *non-stochastic with respect to* Θ.

Kolmogorov's formalization consists not only of this most general definition of the rule of choice but also of the requirement (after Church) that the rule of choice is computable, i.e. the rule of choice is an algorithm.

Definition 3.2. Let us call an (infinite) binary sequence *stochastic* if it is stochastic with respect to any computable rule of choice.

Thus, stochasticness is the second (after lawlessness) candidate to refine the informal concept of randomness. The set of stochastic sequences is large in the measure-theoretic sense (but not in the sense of category). The measure being implicitly present in our consideration is the measure connected with independent trials of tossing a fair coin. Let us now proceed to the consideration of the situations related with the more general class of measures.

Definitions 3.3. (1) *σ-algebras.* A *σ-algebra in* X is a collection of subsets of X which contains X (as an element) and is closed under the two operations: of complementation and of countable union. For any collection $\mathscr{F}$ of subsets of X, there exists a smallest

[4] In the sequel we will say that Man (or the rule of choice) *finds out* the corresponding value.

σ-algebra in X in which $\mathscr{F}$ is included as a subcollection; that σ-algebra is called *the σ-algebra generated by $\mathscr{F}$.*

(2) *Measures and distributions.* Generally speaking, a *measure on X* is a countably additive function whose domain is a σ-algebra in X and whose range is in $[0,+\infty)$. We say that a measure is a *probabilistic measure* or a *probability distribution* if its value on X is equal to 1.

(3) *Measures on Ω.* In the sequel, *a measure on Ω* will be understood in a more restricted sense: it is a probabilistic measure defined on the σ-algebra generated by the collection of all intervals in Ω. So any measure φ on Ω is completely determined by its values on the intervals, that is by the values $\varphi(\Omega_s)$. Let us consider a function f with the domain Ξ: $f(s)=\varphi(\Omega_s)$. This f, related to a measure φ, is called a *quasimeasure.* Often it is convenient to denote a measure and a quasimeasure, correlated one to another, by the same letter. Of course, the notion of a quasimeasure can be defined independently.

(4) *Quasimeasures.* A function $\mu:\Xi\to[0,+\infty)$ is a *quasimeasure on Ξ* (or simply *quasimeasure*) if $\mu(s)=\mu(s\hat{}0)+\mu(s\hat{}1)$ for all s and $\mu(\Lambda)=1$. For any quasimeasure μ let us consider the measure denoted by the same letter μ and defined on the family of all intervals: $\mu(\Omega_s)=\mu(s)$. As $\mu(\Omega)=\mu(\Omega_\Lambda)=\mu(\Lambda)=1$ this measure is a probabilistic measure or a probability distribution.

(5) *Bernoulli measures.* Let $\mathbf{p}=\langle p_1,p_2,\dots,p_k,\dots\rangle$ be a sequence of reals from the segment $[0,1]$, $k\in\mathbb{N}$. Let us define a quasimeasure μ by the equality $\mu(s)=\prod_{1\leqslant k\leqslant n} r_k$, where $s=\langle s_1,\dots,s_n\rangle$, $n=|s|$ and $r_k=p_k$ if $s_k=1$ and $r_k=1-p_k$ if $s_k=0$ (For example, if $p_1=0.3$ and $p_2=0.6$ then $\mu(\langle 1,0\rangle)=0.12$.) The corresponding measure on Ω is called *generalized Bernoulli measure* (or *generalized Bernoulli distribution*) with parameter $\mathbf{p}$. If for all $k\in\mathbb{N}$ it holds $p_k=p$ then this measure is called *Bernoulli measure* (or *Bernoulli distribution*). Finally, if $p=\frac{1}{2}$ then this Bernoulli measure is called the *uniform measure* (or *uniform distribution*). In the sequel the uniform measure will be denoted by the letter η, thus $\eta(\Omega_s)=2^{-|s|}$.

(6) *Computability.* Of special interest is the case of *computable quasimeasure.* The computability of a quasimeasure means that all its values are computable real numbers and the mapping $\mu:\Xi\to[0,+\infty)$ itself is computable. The measure on Ω is called *computable* if it is related to a computable quasimeasure. We will not give here the definition of a computable real number, or the definition of a computable mapping from Ξ into the set of computable real numbers. It is sufficient for our purposes to consider only quasimeasures with rational values. We assume that the notion of a computable mapping (i.e. a mapping determined by an algorithm) from Ξ into the set $\mathbb{Q}$ of rational numbers is clear. Let us mention that for every rational p (more generally, computable real p) the Bernoulli measure with parameter p is computable.

Theorem 3.4. *For the uniform measure: the measure of the set consisting of all the sequences that are non-stochastic with respect to a fixed rule of choice is equal to zero.*

The set of computable rules is countable; therefore we get:

Corollary 3.5. *For the uniform measure: the measure of the set of all stochastic sequences is equal to* 1.

We do not prove now Theorem 3.4 because we will prove below the more general Theorem 3.13. Moreover, we will prove in Section 8 a stronger Theorem 8.7 (and its proof will be simpler than all known proofs of Theorem 3.4).

Theorem 3.4 and Corollary 3.5 are evident particular cases of the analogous theorem and corollary for Bernoulli measures – see Theorem 3.6 and Corollary 3.7 below. The notions of "stochasticness with respect to a given rule" and "stochasticness" that are present in those assertions are defined similarly to the uniform case; the difference is that the constant $\frac{1}{2}$ in the formulation of LLN is replaced by parameter p.

Theorem 3.6. *For a Bernoulli measure: the measure of the set of all sequences that are non-stochastic with respect to a fixed rule of choice is equal to* 0.

Corollary 3.7. *For a Bernoulli measure: the measure of the set of all stochastic sequences is equal to* 1.

In their turn, Theorem 3.6 and Corollary 3.7 are the particular cases of the more general assertions 3.10 and 3.11 for generalized Bernoulli measures. However, we have to define first what is a *stochastic sequence with respect to a generalized Bernoulli measure* on Ω. As we will see the situation is more complicated in this case. In the case of Bernoulli measure we gave the definition of LLN for a sequence and then applied automatically this definition to a subsequence. Now we have to take into account that the probability of appearing of 1 on the seventh (say) place can be different for different subsequences.

To avoid iterated indices let us write $a(m)$ instead of a_m.

Definition 3.8. Let we are given a generalized Bernoulli measure with parameter $\langle p(1), p(2), \ldots, p(m), \ldots\rangle$. A sequence $\mathbf{a} = \langle a(1), a(2), \ldots, a(m), \ldots\rangle$ is called *stochastic with respect to a rule* Θ if the subsequence $\mathbf{b} = \langle a(m_1), a(m_2), \ldots\rangle$ obtained from $\mathbf{a}$ by the application of Θ is finite or satisfies the following condition:

$$\frac{a(m_1) + \cdots + a(m_k)}{k} - \frac{p(m_1) + \cdots + p(m_k)}{k} \to 0$$

as $k \to \infty$. (The sum $a(m_1) + \cdots + a(m_k)$ is equal to the number of ones among the k first terms of the sequence $\mathbf{b}$.)

Definition 3.9. For a generalized Bernoulli measure: a sequence is called *stochastic* if it is stochastic with respect to every computable rule of choice.

Theorem 3.10. *For a generalized Bernoulli measure: the measure of the set of all sequences that are non-stochastic with respect to a fixed rule of choice is equal to* 0.

Corollary 3.11. *For a generalized Bernoulli measure: the measure of the set of all stochastic sequences is equal to* 1.

Now we will prove Theorem 3.10 in the important particular case: the case of the so called quasiuniform measures.

Definition 3.12. A generalized Bernoulli measure with the parameter $\langle p_1, p_2, \ldots, p_m, \ldots \rangle$ is called *quasiuniform* if $p_m \to \frac{1}{2}$ as $m \to \infty$.

Thus we will prove the following theorem:

Theorem 3.13. *For a quasiuniform measure: the measure of all sequences that are non-stochastic with respect to a fixed rule of choice is equal to* 0.

Corollary 3.14. *For a quasiuniform measure: the measure of the set of all stochastic sequences is equal to* 1.

Proof of Theorem 3.13. Let μ be a quasiuniform measure with the parameter $\langle p_1, p_2, \ldots, p_n, \ldots \rangle$, let Θ be a rule of choice. Evidently it suffices to prove the following assertion: if $\varepsilon > 0$ then the μ-measure of the set

$A = \{\mathbf{a} \in \Omega$: $\forall n_0\ \exists n \geqslant n_0$ [the subsequence $\mathbf{b} = \mathbf{b}(\mathbf{a})$ chosen from $\mathbf{a}$ according to the rule Θ has $\geqslant n$ terms and there are at least $(0.5+\varepsilon)n$ ones among its n first terms]$\}$

is equal to 0 (the case when the number of zeros is at least $(0.5+\varepsilon)n$ is analogous).

Let us begin with the remark: the number

$$p = p(\varepsilon) = (0.5+\varepsilon)^{0.5+\varepsilon}(0.5-\varepsilon)^{0.5-\varepsilon}$$

satisfies the inequality $p > 0.5$ if $\varepsilon < 0.5$ (for the proof we can take the logarithm and compute the derivative). Let $p = 0.5 + \delta$, $\delta > 0$.

Lemma 3.13.1. *For $n \in \mathbb{N}$, let B_n be the set of all sequences $s \in \Xi$ of the length n such that there are at least $(0.5+\varepsilon)n$ ones among terms of s. For any n B_n has at most $(1/p)^n$ elements.*

Proof. For a sequence s, let us define $v(s) = (0.5+\varepsilon)^{n(1)}(0.5-\varepsilon)^{n(0)}$ where $n(1)$ and $n(0)$ are the number of ones and the number of zeros in s, respectively. Then $\sum_{s \in B_n} v(s) \leqslant [(0.5+\varepsilon)+(0.5-\varepsilon)]^n = 1$ and on the other hand $v(s) \geqslant p^n$ if $s \in B_n$. These two inequalities yield the inequality $|B_n| \leqslant (1/p^n)$. □

Let $d_n = p_n - 0.5$. We may assume that $|d_n| < \delta/2$ for all n (if this is not the case then we can replace every d_n such that $|d_n| \geqslant \delta/2$ with zero; on all intervals the measure μ is not greater than the new measure μ' multiplied by some constant, and therefore any set having zero μ'-measure has also zero μ-measure).

Thus, let us assume that $\forall n |d_n| < \delta/2$.

Lemma 3.13.2. *For any n and for any $s \in \{0,1\}^n$ the μ-measure of the set $D_s = \{\mathbf{a} \in \Omega$: the subsequence $\mathbf{b}$ chosen from $\mathbf{a}$ by application of the rule Θ has at least n terms and the initial segment of $\mathbf{b}$ with the length n is equal to $s\}$ is not greater than $2^{-n}(1+\delta)^n$.*

Proof. The lemma will be proved by induction on n. The base of induction ($n=0$; in this case $s=\Lambda$ and $D_s=\Omega$) is obvious. The induction step easily follows from the inequality $\mu(D_{s^\frown i}) \leqslant \mu(D_s)(0.5+\delta/2)$, where $s \in \{1,0\}^n$, $i \in \{1,0\}$. Let us prove this inequality.

Let $\mathbf{a} \in D_{s^\frown i}$. Let Man be choosing a subsequence of $\mathbf{a}$ according to Θ. Let $V \subseteq \mathbb{N}$ be the set of numbers of all the terms of $\mathbf{a}$ the value of which Man found out before choosing $(n+1)$th term. Let l be the number of $(n+1)$th chosen term of $\mathbf{a}$ and let $U = V \cup \{l\}$. Let us define $E(\mathbf{a}) = \{\mathbf{a}' \in \Omega$: $a'_k = a_k$ for all $k \in U\}$ and let us define $C(\mathbf{a}) = \{\mathbf{a}' \in \Omega$: $a'_k = a_k$ for all $k \in V\}$. Clearly, $a_l = i$, $E(\mathbf{a}) = \{\mathbf{a}' \in C(\mathbf{a})$: $a'_l = i\}$, and this involves $\mu(E(\mathbf{a})) \leqslant \mu(C(\mathbf{a}))(0.5+\delta/2)$. For fixed s, i we can easily verify that if $\mathbf{a}_1$, $\mathbf{a}_2$ belong to $D_{s^\frown i}$ and $E(\mathbf{a}_1) \neq E(\mathbf{a}_2)$ then $C(\mathbf{a}_1) \cap C(\mathbf{a}_2) = \emptyset$. On the other hand, $D_{s^\frown i} = \bigcup_{\mathbf{a} \in D_{s^\frown i}} E(\mathbf{a})$ and $\bigcup_{\mathbf{a} \in D_{s^\frown i}} C(\mathbf{a}) \subseteq D_s$. Therefore $\mu(D_{s^\frown i}) \leqslant \mu(D_s)(0.5+\delta/2)$. □

Lemmas 3.13.1 and 3.13.2 imply that for any n, $\mu(\bigcup_{s \in B_n} D_s) \leqslant [(1+\delta)/(1+2\delta)]^n$. Obviously A is included in $\bigcup_{n>k} \bigcup_{s \in B_n} D_s$ for any $k \in \mathbb{N}$. As the series $\sum_{n=0}^{\infty} [(1+\delta)/(1+2\delta)]^n$ converges, that inclusion implies that $\mu(A) = 0$. □

3.15. In search of completeness let us give the definition of stochasticness for an arbitrary measure.

Let μ be a measure on Ω (below we will introduce a small restriction to simplify the exposition). As we will see the definition will be rather cumbersome.

First, let us introduce some notations. Let $n_1, \ldots, n_k$ be natural numbers and let $i_1, \ldots, i_k$ belong to $\{0,1\}$. We denote by $A^{n_1,\ldots,n_k}_{i_1,\ldots,i_k}$ the set of all sequences $\mathbf{a}$ from Ω such that $a_{n_1} = i_1$, $a_{n_2} = i_2, \ldots, a_{n_k} = i_k$. (It must be stressed that we do not require that $n_1 < n_2 < \cdots < n_k$). Let us consider the conditional probability of mth term being equal to 1, i.e. the fraction $\mu(A^{n_1,\ldots,n_k,m}_{i_1,\ldots,i_k,1}) / \mu(A^{n_1,\ldots,n_k}_{i_1,\ldots,i_k})$. We denote this fraction by $\mu\left(\begin{smallmatrix} m \\ 1 \end{smallmatrix}\middle|\begin{smallmatrix} n_1, \ldots, n_k \\ i_1, \ldots, i_k \end{smallmatrix}\right)$. To simplify the exposition we will assume that if all n_i are distinct, then the number in the denominator is not zero (it is that restriction which we mentioned above). Let us fix now an arbitrary sequence $\mathbf{a} \in \Omega$ and an arbitrary rule of choice Θ. Our goal is to define the meaning of expression "the sequence $\mathbf{a}$ is stochastic with respect to the rule Θ". The definition will be cumbersome because we must deal not only with the sequence $\mathbf{b}$ obtained from $\mathbf{a}$ by application of the rule Θ (as in the case of Bernoulli measure) and even not only with the way of embedding of $\mathbf{b}$ in $\mathbf{a}$ (as in the case of generalized Bernoulli measure), but also with the history of generating $\mathbf{b}$. The process of generating $\mathbf{b}$ according to the rule Θ can be divided into two steps. In the first step the rule Θ constructs auxiliary sequence $\mathbf{c}$ which consists of all terms of the sequence $\mathbf{a}$ which the rule Θ finds out. In the second step the terms of $\mathbf{b}$ are chosen from $\mathbf{c}$.

More formally: $\mathbf{a}=\langle a(1),a(2),\dots,a(n),\dots\rangle$, $\mathbf{c}=\langle a(n(1)),a(n(2)),\dots,a(n(k)),\dots\rangle$ and the number $n(k)$ is computed, according to Θ, given the tuple $\langle a(n(1)),\dots,a(n(k-1))\rangle$.

Further, the rule Θ given $\langle a(n(1)),\dots,a(n(k-1))\rangle$ (but not given $a(n(k))$) decides whether to include $a(n(k))$ in the final subsequence $\mathbf{b}$. Thus, $\mathbf{b}=\langle a(n(k_1))$, $a(n(k_2)),\dots,a(n(k_j)),\dots\rangle$. As we have mentioned, on both steps application of the rule Θ may give no result, in this case the subsequence $\mathbf{b}$ is finite. In the definition of stochasticness, we demand nothing of the subsequence $\mathbf{b}$ if $\mathbf{b}$ is finite. If $\mathbf{b}$ is infinite then the requirements are as follows:

Let us denote by r_j the conditional probability

$$\mu\left(\begin{array}{c|c} n(k_j) & n(1),n(2),\dots,n(k_j-1) \\ 1 & a(n(1)),a(n(2)),\dots,a(n(k_j-1)) \end{array}\right).$$

Let us consider the difference

$$\delta_j=\frac{r_1+r_2+\cdots+r_j}{j}-\frac{a(n(k_1))+a(n(k_2))+\cdots+a(n(k_j))}{j}.$$

We will say that $\mathbf{b}$ satisfies LLN if $\delta_j\to 0$ as $j\to\infty$. We will say that $\mathbf{a}$ is *stochastic with respect to* the rule Θ if the subsequence $\mathbf{b}$ obtained from $\mathbf{a}$ by application of Θ is finite or satisfies LLN.

Finally, let us call a sequence $\mathbf{a}$ *stochastic* if it is stochastic with respect to any computable rule of choice.

Theorem 3.16. *For a measure μ: the measure of the set of all sequences non-stochastic with respect to a fixed rule of choice is equal to* 0.

Corollary 3.17. *For a measure μ: the measure of the set of all stochastic sequences is equal to* 1.

Definition 3.18. Sequences stochastic (non-stochastic) relative to a measure μ will be called *μ-stochastic* (*μ-non-stochastic*, respectively). And we will allow ourselves to write simply *stochastic* and *non-stochastic* sequences, thus omitting the indication of μ, in the case $\mu=\eta$ (where η is the uniform Bernoulli distribution).

4. The principle of typicalness and the majority principle

As we have seen, the notion of lawfulness is related to the notion of a set of first category, and the notion of stochasticness is related to the notion of a set of measure 1. The lawful sequences are exactly those sequences that belong to nowhere dense sets which are boundaries of effectively open sets (Theorem 2.2). And as we will see, some formalizations of the notion of randomness, not equivalent to stochasticness, also satisfy Corollary 3.5: that is, the set of all sequences which are random according to those formalizations has measure 1. All those formalizations are rather special, however there is a general principle involved in all of them.

That principle is the following one: no event of probability 0 can happen. More precisely: by an event we mean that a sequence obtained by infinite number of coin tossings gets into a fixed set $D \subseteq \Omega$ and by the probability of this event we mean the measure of D; then after Wald [26–28] we postulate that a random sequence belongs to no set of measure 0.

Literal understanding this principle leads to contradiction, as every "one-element" event (that is, the event "to be equal to some fixed sequence") has zero probability. A reasonable refinement is as follows: no *predicted* event of probability 0 can happen. Predictions are expressed in a before fixed language. We will suppose that language is countable (that is, it has countable number of expressions). The set of events definable in that language is therefore countable. Consequently, the union of all definable sets of measure 0 has measure 0, and we forbid a random sequence to belong to that union.

The final version of the principle is as follows.

4.1. *The principle of typicalness.* If a sequence is random then it does not belong to any *definable set of measure* 0. (Cf. [12].)

The mathematical meaning of this principle is that it can serve to check if proposed definitions of randomness are adequate. In every particular case we have only to specify which sets are definable, that is to specify all accepted ways of defining a set. For example, for the concept of stochasticness we accept that a set (of infinite binary sequences) is *definable* if it consists of all sequences non-stochastic with respect to a computable rule of choice (in this case any definable set has measure 0). Thus, the notion of stochasticness – as a possible refinement of the concept of randomness – satisfies the principle of typicalness (due to Theorem 3.16).

4.2. *The majority principle.* Almost all (relative to measure which the considered version of notion of randomness is based on) sequences are random. In other words, the measure of the set of all random sequences is equal to 1.

As it is stated in Corollary 3.17, the notion of stochasticness as a possible refinement of the concept of randomness satisfies the majority principle. As we have seen in Section 2, the notion of lawlessness does not satisfy this principle.

5. Relations between different measures. The principle of distinguishing

An interesting difference between category and measure is that category is uniquely specified by topology of space, whereas the measure $\mu(s) = \mu(\Omega_s)$ of intervals Ω_s, $s \in \Xi$, can be defined in arbitrary way, the only restriction is the equality $\mu(s) = \mu(s\hat{\ }0) + \mu(s\hat{\ }1)$ (where $s\hat{\ }i$ is the extension of s with term i). Quite natural question is: does the notion of typicalness depend on the chosen measure? A possible application of the research in this direction is the identification of an unknown probability distribution by means of tests: the distribution must be such one that the sequence of results of a test is random relative to it.

Let us call two measures *mutually singular*, or *inconsistent*, if there is a set $X \subseteq \Omega$ having measure 0 (i.e. is very small) relative to first measure and having measure 1 (i.e. is very large) relative to second measure. We do not require that if the considered measures are definable in a reasonable sense (for example, are computable) then there is a definable set proving their inconsistency. But it turns out that for different refinements of the notion "definability" the inconsistency can always be proved by means of definable sets. For example, if two mutually singular measures are computable, then the constructive supports of those measures are disjoint [13, Section 35].

Measures that are inconsistent with the uniform measure can be obtained by means of independent tosses of asymmetrical coins. Let $p_n = 0.5 + d_n$ be the probability of 1 in nth tossing ($|d_n| \leqslant 0.5$). We obtain the generalized Bernoulli measure according to Definition 3.3 as follows: $\mu(\Omega_s) = \mu(s) = \prod_{1 \leqslant k \leqslant n} r_k$, where n stands for the length of tuple s and r_k is equal to p_k if kth term of s is equal to 1 and to $(1 - p_k)$ if that term is equal to 0. The measure of an interval Ω_s is equal to the probability of the event "the sequence obtained by tossing of coins belongs to Ω_s".

Clearly, the uniform measure η can be also defined in that way: let us take $p_n = 0.5$ for all n.

Theorem 5.1 (Kakutani [4]). *Let a generalized Bernoulli measure μ has parameter $\mathbf{p} = \langle p_1, p_2, \dots p_n, \dots \rangle$, where $p_n = 0.5 + d_n$ and the series $\sum d_n^2$ diverges. Then the measure μ and the uniform measure η are inconsistent.*

Proof. For every real $\varepsilon > 0$ we will construct a set $X = X(\varepsilon)$ such that $\eta(X) < \varepsilon$ and $\mu(X) > 1 - \varepsilon$. If this is done, then the set

$$Y = \bigcap_n \bigcup_{m \geqslant n} X(2^{-m})$$

will have η-measure 0 and μ-measure 1.

To this end let us consider an auxiliary measure μ' defined in the same way as μ is; namely, we take $d_n' = d_n/2$ for μ'. As

$$(0.5 + d_n/2)^2 = 0.5(0.5 + d_n)\left(1 + \frac{d_n^2}{1 + 2d_n}\right) \geqslant 0.5(0.5 + d_n)\left(1 + \frac{d_n^2}{2}\right),$$

we have

$$\mu'(s)^2 \geqslant \eta(s)\mu(s)\prod_{n=1}^{k}\left(1 + \frac{d_n^2}{2}\right)$$

for every interval Ω_s, where $k = \mathrm{lh}(s)$.

Let us fix a natural $M > \varepsilon^{-2}$. As the series $\sum d_n^2$ diverges we can find N such that for any tuple $s \in I^N$ (where I^N is the set of all binary N-tuples) it holds $\mu'(s)^2 \geqslant \mu(s)\eta(s)M$. Let us denote by S the set of all $s \in I^N$ such that $\mu(s) \geqslant \eta(s)$. We will prove that the set $X = \bigcup_{s \in S} \Omega_s$ can be taken as $X(\varepsilon)$.

First, if $s \in S$ then $\mu'(s)^2 \geqslant M\eta(s)^2$ which involves $\mu'(s) \geqslant \sqrt{M}\eta(s)$ and therefore

$$1 \geqslant \sum_{s \in S} \mu'(s) \geqslant \sqrt{M} \sum_{s \in S} \eta(s).$$

The last inequality implies

$$\eta(X) = \sum_{s \in S} \eta(s) \leqslant M^{-0.5} < \varepsilon.$$

Second, if $s \notin S$ then $\eta(s) \geqslant \mu(s)$ and in the same way we get $\mu'(s) \geqslant \sqrt{M}\mu(s)$, therefore

$$1 \geqslant \sum_{s \notin S} \mu'(s) \geqslant \sqrt{M} \sum_{s \notin S} \mu(s), \qquad \mu(\Omega \backslash X) < \varepsilon$$

and, finally, we get $\mu(X) > 1 - \varepsilon$. □

In a sense two inconsistent measures are like the measure and the category (there is a set of first category having measure 1). Therefore keeping in mind the inconsistency of the notions of lawlessness and typicalness (even if typicalness is restricted to the law of large numbers), we can state one more principle:

5.2. *Principle of distinguishing.* No sequence is random relative to two inconsistent measures.

Theorem 5.3. *Stochasticness as randomness does not satisfy the principle of distinguishing. That is, there are two inconsistent measures and a sequence such that the sequence is stochastic with respect to both measures.*

The proof is based on an easy observation that is stated in the following proposition.

Proposition 5.4. *For an arbitrary quasiuniform measure μ, stochasticness relative to μ is equivalent to stochasticness relative to the uniform measure η.*

Proof. The notations p_m and $p(m)$ was explained in Definitions 3.8 and 3.12 (so $p_m = p(m)$). Since $p_m \to 0.5$ as $m \to \infty$, hence $[p(m_1) + \cdots + p(m_k)]/k \to 0.5$ as $k \to \infty$ for distinct $m_1, \ldots, m_k$. □

Now, to prove Theorem 5.3 it is sufficient to take any quasi-uniform measure μ satisfying conditions of Kakutani's theorem.

So, the second main requirement to a refinement of the notion of randomness – the principle of distinguishing – is not satisfied for the notion of stochasticness. Therefore, in the following sections we will look for more adequate approach.

We conclude this section with two methodological remarks.

Remark 1. Theorem 5.3 can be proved without giving an exact definition of stochasticness relative to the measures implied in the formulation of that theorem. It is significant only to suppose that the definition satisfies the majority principle. For any measure μ

let us denote by S_μ the set of sequences stochastic relative to μ. It suffices now to prove the existence of a measure μ satisfying two following conditions:
(1) μ is inconsistent with η;
(2) $\mu(S_\eta)=1$.

Then, by majority principle we have $\mu(S_\mu)=1$. Therefore the intersection $S_\mu \cap S_\eta$ is non-empty, and any sequence from this intersection contradicts the principle of distinguishing.

Remark 2. Although we reject stochasticness as a *definition* of randomness, stochasticness remains to be one of the most important *applications* of randomness. We would remind the following words of Kolmogorov [5]: "the basis for the applicability of the results of the mathematical theory of probability to real 'random phenomena' must depend on some form of the *frequency concept of probability*, the unavoidable nature of which has been established by von Mises in a spirited manner".

6. Games with finite and infinite sequences

Intuitively speaking, a random sequence is an unpredictable sequence. All games of chance are based on the impossibility of prediction of results of such games. It is not surprising that probability theory arose in the middle of 18th century with observation of such games. In 20th century the nonexistence of winning strategy was taken by von Mises as the base of his approach to probability theory. He wrote:
"This impossibility of affecting the chances of a game by a system of selection, this uselessness of all systems of gambling, is the characteristic and decisive property common to all sequences of observations or mass phenomena which form the proper subject of probability calculus. $\langle\cdots\rangle$ Everybody who has been to Monte Carlo, or who has read descriptions of a gambling bank, knows how many 'absolutely safe' gambling systems, sometimes of an enormously complicated character, have been invented and tried out by gamblers; and new systems are still being suggested every day. The authors of such systems have all, sooner or later, had the sad experience of finding out that no system is able to improve their chances of winning in the long run, i.e., to affect the relative frequencies with which different colours or numbers appear in a sequence selected from the total sequence of the game. This experience forms the total basis of our definition of probability". ([15], the English translation is taken from [16, pp. 24–25]).

Therefore our next trial to formalize the notion of random sequence will be related with the notion of a game. Two players take part in the game – Man and (infinite binary) Sequence. Man tries to guess the values of some terms of Sequence. If he fails then he pays and if he is successful then he gets a prize. Our intuition tells us that a very random Sequence cannot be beaten. Of course we have given a rather vague idea. The notion of a game must be and will be described rigorously.

Let us begin, however, with the description of games in which Man plays with a finite sequence. Finite games will clarify main ideas of the game approach to definition of randomness. In addition, some of obtained facts will be used in the study of infinite sequences.

6.1. Let us fix a number m and a measure φ on the set I^m of all binary m-tuples, or finite binary sequences of length m. We suppose that all values of measure φ are positive rational numbers and that $\varphi(I^m) = 1$. To define such a measure we have to define its value on all one-element sets, i.e. to define a function from I^m into $\mathbb{Q}$. Such a function can be represented as a finite table consisting of constructive objects, therefore it is, in its turn, a constructive object; consequently such a function can be an input and a result of an algorithm (computable function); this fact will be used in the sequel. The measure of a set can be considered as the probability of getting into this set.

Thus, in the definition of a game there are two parameters: the length of tuple m and measure φ on I^m. Such a game is called *m-finite φ-game*.

We will consider two versions of game: "For cash" and "On credit". The rules of game consist of two parts. The first part is common for both versions, the second one is specific.

The common rules are as follows. Let $\mathbf{a} = \langle a_1, a_2, \dots, a_m \rangle$ be a fixed binary m-tuple. The player Man does not know this m-tuple. At the beginning of the game Man has an initial capital V_0, which is a nonnegative rational number. Man makes consecutive moves with numbers from 1 to m. Before kth move Man possesses the capital V_{k-1}, where $k = 1, \dots, m$. The *move* with number k is a pair: a *predicted value* $i(k) \in \{0, 1\}$ and a *bet* $v(k) \in \mathbb{Q}, v(k) \geqslant 0$. After kth move Man finds out true value of a_k. If $i(k) \neq a_k$ (i.e. Man's guess is not correct) Man loses the bet and his capital is decreased by the bet: $V_k = V_{k-1} - v(k)$. If $i(k) = a_k$ (i.e. Man's guess is correct) then his capital is increased by the value of the bet multiplied by a coefficient. This coefficient is defined as follows. Let us denote by I^m_s the set of all m-tuples extending the tuple s. Let us introduce two k-tuples $u = \langle a_1, \dots, a_{k-1}, 1 - i(k) \rangle$ and $w = \langle a_1, \dots, a_{k-1}, i(k) \rangle$. Then (if $i(k) = a_k$)

$$V_k = V_{k-1} + v(k)\varphi(I^m_u)/\varphi(I^m_w).$$

Man's goal is to increase his capital. Therefore, we are interested in the value of *final* capital V_m.

While playing Man can use a strategy, i.e. some instructions how to act in different situations. More precisely, we usually call a *strategy* any mapping from the set of all possible positions into the set of all legal moves. As to the term "position", this is the common term for the formal representation of information known to Man. In our case we could call the position before the kth move the tuple $\langle a_1, \dots, a_{k-1}, V_{k-1} \rangle$. However, we can easily prove (by induction on k) that if we are given a strategy σ in the sense above then in any game for every fixed $V_0, a_1, \dots, a_{k-1}$ we will use the value $\sigma(\langle a_1, \dots, a_{k-1}, V_{k-1} \rangle)$ only for one value of V_{k-1} and this value can be computed given $V_0, a_1, \dots, a_{k-1}$. Therefore we will call a *strategy* any mapping from the set of tuples of the form $\langle a_1, \dots, a_{k-1} \rangle$, $k \leqslant m$ (including the empty tuple) into the set of moves, i.e. the set of pairs $\langle i(k), v(k) \rangle$. If the current move ordered by the strategy is not legal (below we will give the restrictions of value of bet) then the game is over and final capital is equal to the current one.

Now we will describe two versions of the game. Both versions are specified with restrictions of value of bet. In the game "For cash" the restriction is that $v(k) \leqslant V_{k-1}$, thus in this game the inequality $V_k \geqslant 0$ always holds. In the game "On credit" the restriction is that $v(k) \leqslant 1$; thus Man's capital can become negative.

We can easily see that in both versions given a strategy and an initial capital V_0 we can decide effectively whether the strategy makes only legal moves. We will call strategies satisfying this requirement *correct* (for given V_0).

Remark. Informally speaking, the game "On credit" is similar to the logarithm of the game "For cash". If in a m-finite φ-game the measure φ of every tuple is equal to 2^{-m}, then in the game "For cash" the capital after one move can increase at most *twice*. In the game "On credit" after one move the capital can increase at most *by the additive constant* 1. In the game "For cash" it is natural to be interested in the value of the *ratio* V_m/V_0 and in the game "On credit" – in the value of the *difference* $V_m - V_0$. And so on. There is also a logarithmic analogy in the two following theorems.

In the case of finite games (i.e. games with finite sequences) the statements of theorems become easier if we restrict ourselves with *normalized* games, i.e. games with *normalized* initial capital. In the game "For cash" the normalized initial capital is equal to one, in the game "On credit" the normalized initial capital is equal to zero. However, in the sequel in the study of infinite games (i.e. games with infinite sequences) we will use non-normalized games. The reason is that sometimes it is convenient to divide an infinite game into infinite succession of phases, each of them is a finite game. In those finite games initial capital may be non-normalized: it will be equal to the capital in infinite game possessed by Man at the beginning of the phase. All theorems on normalized games can be easily restated for non-normalized games.

It is natural to expect that Man can increase his initial capital if he possesses some information about the sequence which he plays against. The following theorems state this assertion in more precise form.

Theorem 6.1.1. On the game "For cash" (An.A. Muchnik).

1. *Let a measure φ on I^m and a set $S \subseteq I^m$ be given. Then there is a strategy in the normalized m-finite φ-game "For cash" such that for every finite sequence $a \in S$ the strategy guarantees that $V_m \geqslant (\varphi(S))^{-1}$ in the game against a.*

2. *The above mentioned strategy can be effectively found given m, φ and S. That is, there is a computable function f of 4 arguments m, φ, S and x (where $x = \langle x_1, \ldots, x_k \rangle$ is the tuple consisting of known terms of sequence) such that for all m, φ, and S the function $\lambda x.f(m, \varphi, S, x)$ is a strategy satisfying the requirements of item 1.*

3. *The bound from item 1 cannot be improved in the following sense:* $\forall m \, \forall \varphi \, \forall S \neq \emptyset \, \forall \sigma \, \exists a \in S$ *[in the normalized m-finite φ-game "For cash", if Man plays according to the strategy σ against a then $V_m \leqslant (\varphi(S))^{-1}$].*

Theorem 6.1.2. On the game "On credit" (An.A. Muchnik).

1. *Let a measure φ on I^m, a subset $S \subseteq I^m$ and a rational number $r < -\ln(\varphi(S))$ be given. Then there is a strategy which for every finite sequence $a \in S$ guarantees the inequality $V_m > r$ in the normalized m-finite φ-game "On credit" against a.*

2. *The strategy of item 1 can be effectively found given m, φ, S and r. That is, there is a computable function f of 5 arguments m, φ, S, r and x such that $\forall m \forall \varphi \forall S \forall r\, \lambda x.f(m, \varphi, S, r, x)$ is a strategy satisfying the requirements of item 1.*

3. *The bound from item 1 cannot be improved in the following sense. Let rationals ρ and r be fixed and let $0 < \rho \leqslant 1$, $r > -\ln(\rho)$. Then $\exists m_0 \forall m \geqslant m_0 \exists \varphi \exists a \in I^m [(\varphi(\{a\}) < \rho)$ & $\forall \sigma$ [in the normalized m-finite φ-game "On credit" against a, if Man plays according to the strategy σ then $V_m < r$]].*

Let us call the *uniform* measure on I^m the measure φ defined by equality $\varphi(\{a\}) = 2^{-m}$ for every $a \in I^m$.

Theorem 6.1.3. On the game "On credit" for the uniform measure (An.A. Muchnik).

1. *Let a subset $S \subseteq I^m$ and a rational $r < -lb(\varphi(S))$ be given, where φ is the uniform measure and* lb *is logarithm to the base 2. Then there is a strategy in the normalized m-finite φ-game "On credit" such that if Man plays according to this strategy against any sequence $a \in S$ then the inequality $V_m > r$ holds.*

2. *The strategy of item 1 can be effectively found given m, S and r.*

3. *The bound from item 1 cannot be improved in the following sense*: $\forall n \forall m \geqslant n \forall s (|s| = n) \forall \sigma \exists a \in I^m_s$ [*if in the normalized m-finite φ-game "On credit" Man plays against a according to the strategy σ then $V_m \leqslant -\mathrm{lb}(\varphi(I^m_s)) = n$*].

Remark 6.1.4. The three last theorems can be generalized to non-normalized games. In the first theorem we have to replace V_m with the ratio V_m/V_0, in the last two theorems we have to replace V_m with the difference $V_m - V_0$. The statements on non-normalized games are trivial consequences of the corresponding statements on normalized games.

Some remarks on the proofs:

1. In this paper we will prove only the first items of Theorems 6.1.1 and 6.1.3. These proofs will be used in basic Theorems 9.1 and 9.2.

2. The second items of all theorems follow from the effectiveness of the proofs of the first items.

Now we are going to prove the first item of theorem on the game "For cash". Let us prove first the following lemma.

Lemma 6.1.5. *Let φ be a measure on I^m and let V be a nonnegative function from $I^{\leqslant m}$ ($I^{\leqslant m}$ stands for the set of all k-tuples with $k \leqslant m$) into $\mathbb{Q}$ such that $V(\Lambda) = 1$ and for any tuple s if $|s| < m$ then the following holds*:

$$V(s)\varphi(s) = V(s\hat{\ }0)\varphi(s\hat{\ }0) + V(s\hat{\ }1)\varphi(s\hat{\ }1) \qquad (*)$$

(here $\varphi(s)$ denotes $\varphi(I_s^m)$, and I_s^m is the set of all tuples from I^m that extend s). *Then there is a correct strategy σ in the normalized m-finite φ-game "For cash" such that for any k-tuple s with $k \leqslant m$ if Man plays according to σ against s then his capital after k moves is equal to $V(s)$.*

Proof. Let us describe the strategy σ. Assume that in position s ($|s| = k-1$) the current capital is equal to $V(s)$. Man has to predict the kth term. To this end Man determines the least of two numbers $V(s\hat{}i)$, $i \in \{0,1\}$ (if they are equal then Man chooses any of them); without loss of generality we can assume that this number is $V(s\hat{}1)$. Then $V(s\hat{}1) \leqslant V(s)$ (this follows from $(*)$, as φ is a measure). Man bets the capital $v = V(s) - V(s\hat{}1)$ on 0. Then after this move the capital in the game against $s\hat{}1$ will be equal to $V(s) - v = V(s\hat{}1)$ and in the game against $s\hat{}0$ it will be equal to

$$V(s) + v\varphi(s\hat{}1)/\varphi(s\hat{}0) = [V(s)\varphi(s) - V(s\hat{}1)\varphi(s\hat{}1)]/\varphi(s\hat{}0),$$

which is equal to $V(s\hat{}0)$ because of $(*)$. □

6.1.6. *Proof of item* 1 *of Theorem* 6.1.1. Let us put $V(s) = (\varphi(S))^{-1}$ if $s \in S$ and $V(s) = 0$ if $s \in I^m \setminus S$. Then $\sum_{s \in I^m} V(s)\varphi(s) = 1$. Further, let us define

$$V(t) = (\varphi(t))^{-1}[V(t\hat{}0)\varphi(t\hat{}0) + V(t\hat{}1)\varphi(t\hat{}1)]$$

for $|t| < m$ (here we use induction on $m - |t|$). By induction we can prove that for all $l \leqslant m$, the equality $\sum_{s \in I^l} V(s)\varphi(s) = 1$ holds. In particular, for $l = 0$ when I^l consists of the single empty sequence Λ, we get $V(\Lambda) = 1$ as $\varphi(\Lambda) = 1$.

On the other hand, function V satisfies the equality $(*)$ from Lemma 6.1.5. Applying Lemma 6.1.5 we get the desired strategy. □

Remark 6.1.7. In the sequel we will consider finite games "For cash" – the fragments of infinite games – in which the measure of I^m is not equal to 1. Let the measure of I^m be equal to $\varphi_0 \leqslant 1$. Similarly to the proof of item 1 of Theorem 6.1.1 we can easily show (this is left to the reader) that in this case we can guarantee the inequality $V_m \geqslant (\varphi(S))^{-1}\varphi_0$ – in the conditions of item 1 of Theorem 6.1.1.

6.1.8. *Proof of item* 1 *of Theorem* 6.1.3. Let us consider function f defined on $I^{\leqslant m}$ by equality $f(t) = \mathrm{lb}(\varphi(S \cap I_t^m)/\varphi(I_t^m)) - \mathrm{lb}(\varphi(S))$. (If $S \cap I_t^m = \emptyset$, then $f(t) = -\infty$). Now we can define a strategy, which we are looking for. If $f(t\hat{}0) \geqslant f(t)$ then the strategy predicts zero as value of $(|t|+1)$th term, and the bet is equal to $f(t\hat{}0) - f(t)$. If $f(t\hat{}1) \geqslant f(t)$ then the strategy predicts one as value of $(|t|+1)$th term, and the bet is equal to $f(t\hat{}1) - f(t)$. It is clear that $f(\Lambda) = 0$ and largest increase of current capital after one move is 1. For the strategy just described, we claim that for every $t \in I^{\leqslant m}$ the current capital is not less than $f(t)$. It is sufficient to prove that $\forall t\, f(t) \geqslant (f(t\hat{}0) + f(t\hat{}1))/2$. Due to the definition of function f, the last inequality is deduced from the next one: $\sqrt{xy} \leqslant (x+y)/2$, where $x = \varphi(S \cap I_{t\hat{}0}^m)$ and $y = \varphi(S \cap I_{t\hat{}1}^m)$.

Our strategy does not satisfy the requirement of rationality of bets. To correct this defect, we can define a new strategy whose bets will be rational but very close to the corresponding bets of previous strategy. If $t \in S$, then $f(t) = -\mathrm{lb}(\varphi(S))$. So the new strategy has capital $V_m > r$ when $t \in S$. □

6.2. Let us turn now to infinite games, i.e. games in which the active player – Man – plays against an infinite sequence $\mathbf{a} \in \Omega$. Additional parameters are a fixed quasimeasure μ on Ξ (below we will usually assume that μ is computable) and a positive rational which is Man's initial capital. According to 3.3.4, that quasimeasure induces the corresponding probabilistic measure, or probability distribution, μ on Ω.

The game, called *μ-game* (or simply *game*, if it is clear what a measure is considered), runs as follows. Man predicts values (unknown to him) of some terms a_n of the sequence $\mathbf{a}$ and bets on predicted value a capital not greater than his current capital (so the game is "For cash"). If Man fails then Man's capital is decreased by the value of his bet and if he is successful then his capital increases by a value depending on his bet and on the a priori probability of the predicted event, that probability being determined making use of μ. Man's prediction is based on the true values of already predicted (correctly or incorrectly) terms – it is the only information that Man gets, and Man's goal is not to increase the number of correct predictions, his goal is to increase his capital. The game is for cash therefore every bet cannot exceed the current capital. There are three distinctions from above considered game "For cash": first, Man plays against infinite sequence; second, the predicted terms may be chosen in any order (not only in succession); third, the current move can be undefined.

More precisely, the kth ($k = 1,2,3,\dots$) *move* in μ-game consists in the following. Man chooses a *number* $n = n(k)$ of the term a_n being predicted, its *predicted value* $i = i(k) \in \{0,1\}$, and a bet $v = v(k)$, which is a nonnegative rational not greater than the value V_{k-1} of Man's capital before kth move. (Let us recall that V_0 is the initial capital.) Necessary requirement: $n(k)$ cannot coincide with any $n(l), l < k$, i.e. repeating predictions are impossible. Thus, every move is a triple of numbers $\langle n, i, v\rangle$.

Man's capital V_k *after* the kth move is defined as follows. If Man's guess is incorrect, i.e. $a_n \neq i$, then $V_k = V_{k-1} - v(k)$ (the bet is lost).

Let us assume that Man's guess is correct, i.e. $a_{n(k)} = i(k)$. We define

$$A = A(k-1) = \{\mathbf{a}' \in \Omega : a'_{n(l)} = a_{n(l)} \quad \text{for } l = 1,2,\dots,k-1\},$$
$$A_i = A_i(k) = \{\mathbf{a}' \in A : a'_{n(k)} = i\} \quad \text{for } i = 0,1.$$

Then

$$V_k = V_{k-1} + v(k)\mu(A_{1-i(k)})/\mu(A_{i(k)}).$$

This definition is due to the requirement of "fairness" of μ-game: the mathematical expectation of gain (= change of capital) in one move must be equal to zero. We will explain later what to do if $\mu(A_{i(k)}) = 0$.

After computing V_k Man either makes the $(k+1)$th move, or refuses to do the $(k+1)$th move. If Man never refuses to make current move then the game lasts infinitely long. If Man refuses to make the kth move then the game stops and all values $n(l)$, $i(l)$, $v(l)$ with $l \geqslant k$ are undefined. In this case $V_l = V_{k-1}$ for all $l \geqslant k$.

Now we are going to discuss the case $\mu(A_{i(k)}) = 0$. Of course it may not happen if $\mu = \eta$, where η is the uniform measure. But generally speaking it is possible. If we do not want to impose certain restrictions on the measure μ we must add two complementary rules to the list of game rules:

Complementary Rule 1. If $\mu(A(k)) = 0$ then the game stops after the kth move. Since $\mu(\Omega) = 1$, the first move can be made in every game. As $A_0(k) \cup A_1(k) = A(k-1)$, the rule secures that if the kth move is made then at least one of the sets $A_0(k)$ and $A_1(k)$ has positive measure. So the $(k+1)$th move *can* be made in the case $\mu(A(k)) \neq 0$.

Complementary Rule 2. Suppose that the kth prediction is correct (i.e. $a_{n(k)} = i(k)$) and $\mu(A_{i(k)}) = 0$. Then Man's gain is declared to be equal to plus infinity. In this case the game stops after the kth move and we define $V_k = +\infty$.

The consequences of stopping the game by reason of zero measure are the same as those of Man's refusal to make the move next in turn.

Thus, for any game the function V_k (as a function of k) is defined on the whole set $\mathbb{N}$ of natural numbers, whereas the functions $n(k), i(k)$ and $v(k)$ are defined on an initial segment of $\mathbb{N}$ (that segment may coincide with $\mathbb{N}$).

Definition 6.2.1. The result of μ-game: we define that Man *wins* if $\sup_k V_k = +\infty$ (Of course, if it occurs that $V_k = +\infty$ for some k then Man wins.)

This definition has a reasonable sense. Assume that sequence **a** is the sequence of moves of the second player – let us call him Casino – who has also an initial capital (unknown to Man as well as the sequence **a** is); and all pay-offs to Man are paid from the capital of Casino, and vice versa every loss of Man is added to the capital of Casino. Then the final win of Man means that Casino is brought to ruin, however large its initial capital is.

As well as in the case of finite games, it is convenient to use the notion of strategy in an infinite game. As we know, a strategy is a way of playing, i.e. an instruction which determines the current move of player given current position. However, because of the same reason as in the case of finite games, it is natural to include in the domain of strategy not all information known to Man (the numbers and values of known terms of the sequence, the order in which they were predicted, the current capital) but only the tuple of known values in the order in which they were predicted.

Definitions 6.2.1.1. A *strategy* is a mapping from the set of all tuples of the form $\langle \alpha_1, \alpha_2, \ldots, \alpha_{k-1} \rangle$, $k = 1, 2, 3, 4, \ldots; \alpha_l = 0, 1$ (the meaning is: $\alpha_l = a_{n(l)}$) into the set of moves, i.e. into the set of triples $\langle n, i, v \rangle \in \mathbb{N} \times \{0, 1\} \times \mathbb{Q}$ (the meaning is: $n = n(k), i = i(k), v = v(k)$ – see the above description of kth move). If this mapping is computable then we call the strategy *computable*. A *winning* strategy is a strategy such that if Man with the initial capital $V_0 = 1$ plays according to the strategy then he wins.

Remark 1. We do not suppose a strategy to be *totally defined* and *correct*. This means that it may happen that Man cannot make the kth move, next in turn, according to the strategy because it is undefined or is not legal. In this case Man refuses to make a move, and as we have said, the values $n(k)$, $i(k)$ and $v(k)$ are considered to be undefined and the capital does not change: $V_{k-1} = V_k = V_{k+1} = V_{k+2} = \cdots$. It is clear that in such a game Man loses (provided that $V_{k-1} \neq +\infty$).

Remark 2. We do not require every term a_k of sequence $\mathbf{a}$ to be ever predicted (correctly or not).

Remark 3. The order of predictions, as well as predicted values and bets, depends in general case not only on measure μ but also on the sequence which Man plays against. To emphasize this dependence we will (when necessary) write $n(\mathbf{a},k)$, $i(\mathbf{a},k)$, $V_k(\mathbf{a})$, $v(\mathbf{a},k)$ instead of $n(k)$, $i(k)$, V_k, $v(k)$. This notation has exact meaning if μ, V_0, and a strategy are given.

Definition 6.2.1.2. Let a quasimeasure μ, an initial capital V_0 and a strategy σ be fixed. For any finite sequence $s = \langle s_1, \ldots, s_m \rangle \in \Xi$ let us define the set

$$A_s = \{\mathbf{a} \in \Omega\colon\ a_{n(\mathbf{a},l)} = s_l \quad \text{for all } l = 1, 2, \ldots, m\}.$$

In particular, $A_\Lambda = \Omega$ (recall that Λ denotes the empty sequence). We can easily see that if $\mathbf{a}, \mathbf{b} \in A_s$ then $V_m(\mathbf{a}) = V_m(\mathbf{b})$. We denote by $V(s)$ the value of the function V_m, where $m = \mathrm{lh}(s)$, on any $\mathbf{a} \in A_s$ (if A_s is empty then $V(s)$ is undefined).

We will say that a finite set $S \subset \Xi$ *gives a partition* if

(i) $\bigcup_{s\in S} \Omega_s = \Omega$;

(ii) for all different p, q from S the intervals Ω_p and Ω_q are disjoint; and

(iii) for all $s \in S$, $\mu(A_s) \neq 0$.

The following proposition gives a global consequence of the (local) property from the rule of μ-game stating that the mathematical expectation of gain for one move is equal to zero.

Proposition 6.2.2. *Let a set S give a partition. Then the equality $\sum_{s\in S} \mu(A_s)V(s) = V_0$ holds.*

Proof. We will prove the equality by induction on the sum n of the lengths of all elements $s \in S$. If $n = 0$ then the equality is obvious. Let the equality holds for all $n < m = \sum_{p\in S} \mathrm{lh}(p)$. Let $s \in S$ be one of the longest tuple from S. We denote $k = \mathrm{lh}(s)$. Let, for the sake of definitiveness, $s = u\hat{\ }1$. From the definition of s and the properties of S we can easily deduce that $s' = u\hat{\ }0$ also belongs to S. Evidently $(S \setminus \{s, s'\}) \cup \{u\}$ is partition and it is sufficient to prove

$$V(s)\mu(A_s) + V(s')\mu(A_{s'}) = V(u)\mu(A_u). \qquad (*)$$

For the sake of definitiveness, let us assume that $i(\mathbf{a},k)=0$ for some (hence for every) sequence $\mathbf{a}\in A_u$. Then by definition,

$$V(u\hat{\ }1)=V(u)-v \quad (\text{where } v=v(\mathbf{a},k)),$$

$$V(u\hat{\ }0)=V(u)+v\mu(A_{u\hat{\ }1})/\mu(A_{u\hat{\ }0}),$$

and the equality $(*)$ becomes obvious. □

The function of capital has some additional properties in the case of continuous strategies. Since this place we will suppose that the initial capital V_0 is equal to 1.

Definition 6.2.3. Let us call a strategy σ *continuous* for μ-game if it is totally defined, correct and, playing μ-game against any sequence $\mathbf{a}=\langle a_1,a_2,\ldots\rangle$, for every k it predicts on kth move the value of kth term a_k.

If a strategy σ is continuous then in the game according to σ against any finite sequence $s\in\Xi$ of length k Man makes exactly k moves; this game coincides with above considered finite game "For cash". We denote by $V(s)=V^\sigma(s)$ the final Man's capital in the game against s (i.e. after k moves); it is clear that $V(\Lambda)=1$. It is also clear that $V^\sigma(s)=V_k(\mathbf{a})$ for any infinite sequence $\mathbf{a}$ extending s. The function $V^\sigma(s)$ will be called *capital function* (attached to the strategy σ).

Proposition 6.2.4. *For any μ-game a continuous strategy σ and the capital function V^σ attached to σ either both are computable, or both are uncomputable.*

Proof. Obvious as in every μ-game not only σ determines the function V^σ, but also, in the case of continuous strategies, V^σ uniquely determines σ. (Let us remind our assumptions of computability of the quasimeasure μ.) □

Proposition 6.2.5. *For a fixed measure μ: a function V from Ξ into the set of nonnegative rationals is the capital function in μ-game attached to some continuous strategy iff the product $V(s)\mu(s)$ is a quasimeasure. (Recall that $\mu(s)=\mu(\Omega_s)$).*

Proof. Let V be a capital function and $\mathrm{lh}(s)=k-1$. In position s Man predicts the value of kth term. Without loss of generality we may assume that Man bets some capital v on the value $a_k=0$ (the case of the bet on $a_k=1$ is entirely symmetrical). Then

$$V(s\hat{\ }1)=V(s)-v, \quad V(s\hat{\ }0)=V(s)+v\mu(s\hat{\ }1)/\mu(s\hat{\ }0),$$

which involves $V(s)\mu(s)=V(s\hat{\ }0)\mu(s\hat{\ }0)+V(s\hat{\ }1)\mu(s\hat{\ }1)$, i.e. $V(s)\mu(s)$ is a quasimeasure. The converse assertion in fact was proved in Lemma 6.1.5. □

Now we will give a generalization of the notion of continuous strategy – the notion of monotone strategy.

Definition 6.2.6. A strategy (not necessary totally defined) is called *monotone* if in the game according to it the numbers of predicted terms increase (i.e. $n(k+1)>n(k)$).

The requirement of monotony is weaker than the requirement of continuity; however the possibilities of Man are enlarged rather slightly; the following proposition formalizes the meaning of this remark.

Proposition 6.2.7. *If Man has a totally defined and correct monotone computable winning strategy then he has also a continuous computable winning strategy.*

Proof. Man has to predict all terms between $n(k)$ and $n(k+1)$ betting zero on arbitrary value. □

As we will show below, the requirement of monotony restricts very much the power of strategies (see Theorems 9.1 and 9.5). Continuous strategies play a very important role in our analysis. In particular, the action of some non-continuous strategies will consist of continuous phases, i.e. of finite games (we have already told about this point). This can be illustrated by the proofs of Theorems 9.1 and 9.2; in those proofs the row $\mathbb{N}$ of natural numbers is partitioned into *zones*, and for every zone there is corresponding continuous strategy which acts within that zone.

7. Predictable and unpredictable sequences

Now we turn to the keynote definition directed towards the use of games for mathematical analysis of the notion of randomness. The idea is as follows: if Man has enough information about terms of an infinite sequence to win in the game against this sequence is not random.

Definition 7.1. If σ is a winning strategy in μ-game against a sequence $\mathbf{a}$ then $\mathbf{a}$ is called *μ-predictable via σ*. A sequence $\mathbf{a}$ is called *μ-predictable* if there is a computable winning strategy in μ-game against $\mathbf{a}$; in other case the sequence is called μ-unpredictable. Let us recall that we suppose the initial capital V_0 to be equal to 1: this requirement is included into Definition 6.2.1.1. However, it is evident that if $\mathbf{a}$ is μ-predictable then for every rational $V_0>0$ there is a winning (in a naturally generalized sense) strategy in μ-game against $\mathbf{a}$ with initial capital V_0.

Let us look if the notion of μ-unpredictable, as "model" of the notion of randomness, satisfies two criteria stated in Sections 4 and 5.

Theorem 7.2 (theorem of typicalness). *The set of all infinite sequences* $\mathbf{a}$ *that are μ-predictable via a strategy σ has μ-measure* 0.

This theorem, which is evidently similar to Theorem 3.16 (and even stronger than Theorem 3.16, as we will see in the proof of Theorem 7.4), will be deduced in the next section (see 8.10) from another theorem of the same kind, namely, from relativized versions of Theorems 8.7 and 8.9.

Corollary 7.3 (majority theorem). *The set of all μ-unpredictable sequences has μ-measure* 1.

Before studying the question whether the notion of unpredictability satisfies our second main principle – the principle of distinguishing – we will consider relations between unpredictability and our previous notion – stochasticness.

Theorem 7.4. *All η-unpredictable sequences are η-stochastic.* (Recall that η denotes the uniform measure.)

(This theorem can be proved also for other measures, however we will not expose the proof here.)

Proof. Let us assume that a sequence **a** is not η-stochastic, i.e. there is a computable rule which chooses from **a** a subsequence **b** not satisfying the law of large numbers. Without loss of generality we may assume that in **b** there occur ones more than zeros; more precisely, there is a rational $p = 0.5 + \delta > 0.5$ such that there are infinitely many m such that there are at least pm ones among m first terms of **b**.

The winning strategy in η-game against **a** is as follows.

First, the predicted terms are chosen in the same order as the computable rule of choice proving nonstochasticness of **a** does.

Second, if this rule does not include a current term in the subsequence **b** then Man bets zero on this term.

Third, if a current term is included in the subsequence **b** then Man bets $V\delta$ on 1, where V is his current capital.

According to our agreement the initial Man's capital V_0 is equal to 1. And it is clear that after choosing n terms of the sequence **b** the current Man's capital is equal to

$$V_n = (1+\delta)^{n_1}(1-\delta)^{n_0} = [(1+\delta)^{\alpha(n)}(1-\delta)^{1-\alpha(n)}]^n,$$

where n_1 and n_0 are the numbers of ones and zeros, respectively, among n first terms of sequence **b**, $\alpha(n) = n_1/n$. Thus, if n is such that $n_1 \geqslant np$ then

$$V_n \geqslant [(1+\delta)^{0.5+\delta}(1-\delta)^{0.5-\delta}]^n.$$

But $(1+\delta)^{0.5+\delta}(1-\delta)^{0.5-\delta} > 1$ for $0 < \delta \leqslant 0.5$ (this fact can be easily verified by means of taking logarithm of the function $\lambda x.\ (1+x)^{0.5+\delta}(1-x)^{0.5-\delta}$ and analyzing the derivative). Hence, as the inequality $n_1 \geqslant np$ holds for infinitely many n and δ does not depend on n, the function V_n is unbounded. $\square$

Thus, all unpredictable sequences are stochastic. The strategy constructed in Theorem 7.4 depends on the rational p and on that element from $\{0, 1\}$ for which the amount of its occurences in many beginnings of the sequence is greater than the amount of the other element's occurences. But we can prove the following fact: *for every place selection rule, there is a strategy (computable if the rule is computable) such that any sequence which is non-stochastic via this rule is predictable via this strategy.*

Now we want to prove that the class of stochastic sequences is strictly greater than the class of unpredictable sequences. This result will be proved not by presenting an example of stochastic predictable sequence (as it is usually done in similar cases), but in another way.

At first we will prove:

Theorem 7.5 (theorem of distinguishing). *If measures μ and λ are computable and inconsistent then no sequence is μ-unpredictable and λ-unpredictable simultaneously.*

Corollary 7.5.1. *If measures μ and λ are computable and inconsistent then the set of all λ-unpredictable sequences has μ-measure* 0.

To prove the corollary we have to recall Corollary 7.3.

Proof of Theorem 7.5. The following lemma is true.

Lemma 7.5.2. *For every $\varepsilon>0$ there are a natural $k=k(\varepsilon)$ and two sets $S(\varepsilon)$, $T(\varepsilon)$ such that $S(\varepsilon)$ and $T(\varepsilon)$ are disjoint, their union equals to I^k, and $\mu(X(\varepsilon))<\varepsilon$, $\lambda(Y(\varepsilon))<\varepsilon$, where*

$$X(\varepsilon)=\bigcup_{s\in S(\varepsilon)}\Omega_s,\qquad Y(\varepsilon)=\bigcup_{t\in T(\varepsilon)}\Omega_t.$$

Proof. Because of inconsistency there is a set $X\subseteq\Omega$ such that $\mu(X)=0$ and $\lambda(Y)=0$, where $Y=\Omega\backslash X$. There are two open sets $G_X,G_Y\subseteq\Omega$ covering X and Y, respectively. (therefore $G_X\cup G_Y=\Omega$) and such that $\mu(G_X)<\varepsilon$, $\lambda(G_Y)<\varepsilon$. As Ω is compact, there are a natural k and disjoint sets $S(\varepsilon),T(\varepsilon)$ such that $S(\varepsilon)\cup T(\varepsilon)=I^k$ and

$$X(\varepsilon)=\bigcup_{s\in S(\varepsilon)}\Omega_s\subseteq G_X,\qquad Y(\varepsilon)=\bigcup_{t\in T(\varepsilon)}\Omega_t\subseteq G_Y.\qquad\square$$

Let us return to the proof of Theorem 7.5. Note that because of computability of the measures, given a rational $\varepsilon>0$ we can compute by exhaustion a number $k(\varepsilon)$ and sets $S(\varepsilon),T(\varepsilon)\subseteq I^{k(\varepsilon)}$ satisfying conditions from the lemma.

Let us construct by induction an increasing sequence of natural $d(1), d(2), d(3),\ldots$ such that $d(1)=0$, and for every m: if

$$\delta_m=\min\{\mu(\Omega_s),\ \lambda(\Omega_s)\colon s\in I^{d(m)},\mu(\Omega_s)>0,\ \lambda(\Omega_s)>0\},\qquad \varepsilon_m=\delta_m 2^{-2m-1}$$

then $d(m+1)=k(\varepsilon_m)$ (k is defined in the lemma).

Let us denote by $\mathbf{a}$ an infinite sequence which Man plays against. We will construct two *continuous* strategies σ^μ and σ^λ. Man wins in μ-game via σ^μ or Man wins in λ-game via σ^λ.

Let us divide the infinite game in the sequence of m-fragments. In *the game on the mth zone* Man predicts the values of terms of the sequence $\mathbf{a}$ with numbers $d(m)+1, d(m)+2, \ldots, d(m+1)$. There are two strategies σ_m^μ and σ_m^λ with symmetrical qualities (1) and (2):

(1) if $\mathbf{a} \in X(\varepsilon_m)$ then σ_m^μ increases the captial $V_{d(m)}$ at least 2^{2m} times and σ_m^λ decreases $V_{d(m)}$ at most twice during the game on the mth zone,

(2) if $\mathbf{a} \in Y(\varepsilon_m)$ then σ_m^λ increases the captial $V_{d(m)}$ at least 2^{2m} times and σ_m^μ decreases $V_{d(m)}$ at most twice during the game on the mth zone.

(X, Y are defined in the lemma.)

The strategy σ_m^λ is obtained by application of Theorem 6.1.1 and Remarks 6.1.4, 6.1.7 to the finite game on the mth zone. This game is against the sequence $\langle \mathbf{a}(d(m)+1), \mathbf{a}(d(m)+2), \ldots, \mathbf{a}(d(m+1))\rangle$, with the measure $\varphi(s) = \lambda(s_0\hat{\ }s)$, where $s_0 = \langle \mathbf{a}(1), \mathbf{a}(2), \ldots, \mathbf{a}(d(m))\rangle$, with the initial captial $V_{d(m)}/2$ and with the set $S = \{s\colon s_0\hat{\ }s \in T(\varepsilon_m)\}$ ($T(\varepsilon)$ is defined in the lemma). The strategy σ_m^μ is constructed analogously. The captial $V_{d(m)}$ cannot decrease more than twice, because only half of $V_{d(m)}$ "takes part" in the game on the mth zone. The "increasing" quality of σ_m^μ, σ_m^λ takes place owing to the construction of $d(m+1)$.

The strategy σ^μ is the union of the strategies σ_m^μ (for λ symmetrically). As $\forall \varepsilon > 0$ $X(\varepsilon) \cup Y(\varepsilon) = \Omega$, at least one of two sets

$$M_\mu = \{m\colon \mathbf{a} \in X(\varepsilon_m)\}, \qquad M_\lambda = \{m\colon \mathbf{a} \in Y(\varepsilon_m)\}$$

is infinite. If $|M_\mu| = \infty$ then σ^μ wins in μ-game, if $|M_\lambda| = \infty$ then σ^λ wins in λ-game. Evidently the strategies σ^μ, σ^λ are computable and monotone. □

Definition 7.6. A sequence is called *monotonically μ-predictable* if there is a monotone computable winning strategy in the μ-game against this sequence; if this is not the case then the sequence is called *monotonically μ-unpredictable.*

Remark 7.6.1 (*on Definition* 7.6). As we will see below (Corollary 9.5.1), in the general case η-predictability does not imply monotonic η-predictability.

From the proof of Theorem 7.5 we obtain the following fact:

Proposition 7.6.2. *If two measures are computable and inconsistent then every sequence is monotonically predictable for one of these measures.*

Earlier (before Theorem 7.5) we intended to give an implicit construction of a stochastic predictable sequence. But now we can state even monotonic predictability of the sequence.

Corollary 7.7 (from Theorem 7.5). *There are η-stochastic monotonically η-predictable sequences.* (Recall that η denotes the uniform measure.)

Proof. There is a computable quasiuniform measure satisfying conditions of Theorem 5.1 (it is sufficient to set $d_n = 0.5n^{-0.5}$); let us denote it by μ. Due to Proposition 5.4 the set of all η-stochastic sequences coincides with the set of all μ-stochastic sequences and therefore it has μ-measure 1 (due to Corollary 3.14). At the same time, the set of all monotonically η-unpredictable sequences has μ-measure 0 due to Corollary 7.5.1 reformulated for monotonic λ-unpredictability. (We can get this strengthening due to monotonicity of both strategies constructed in the proof of Theorem 7.5.) □

Definition 7.8. According to our general agreement (on omitting of the indication of uniformness of measure), an η-predictable (η-unpredictable) sequence will be called also simply *predictable* (*unpredictable*).

In conclusion of this section we would state that Theorems 7.5 and 7.2 give hope that the notion of unpredictable sequence is rather adequate definition of the notion of randomness. However, there is another adequate definition, which we turn now to.

8. Chaotic sequences

We have to begin with some definitions important for algorithmic probability theory.

Definition 8.1. A nonnegative function defined on Ξ is called *semimeasure on* Ξ (or simply *semimeasure*) if it is not greater than some quasimeasure on Ξ (see Definitions 3.3.4).

Our notion of semimeasure differs from the notion of semimeasure given in [2, 23] (English translation [24]); however assertions on semimeasures in the sense of those papers are valid for semimeasures in our sense; they can be proved similarly.

We are particularly interested in semimeasures satisfying the following definition.

Definition 8.2. A semimeasure γ is called *recursively enumerable from below* (briefly, r.e.b.) if the set ("undergraph")

$$\{\langle s,q\rangle \in \Xi \times \mathbb{Q}\colon\ q < \gamma(s)\}$$

is recursively enumerable ($\mathbb{Q}$ stands for the set of rational numbers).

Why we have to consider only such semimeasures? We will interpret $-\mathrm{lb}(\gamma(s))$ (where $\mathrm{lb}(x)$ is the logarithm of x to the base 2) as the quantity of information which is required for the complete description of the object s. Once found, a description is preserved; however, it is possible that later a shorter description will be found. Thus, $-\mathrm{lb}(\gamma(s))$ can be approximized from above and $\gamma(s)$ approximized from below. This is the source of the definition.

On the family of all semimeasures the following quasiorder is defined:

$\gamma_1 \leqslant \gamma_2$ means $\exists c\ \forall s[\gamma_1(s) \leqslant c\gamma_2(s)]$.

Theorem 8.3 (Gacs, [2, Theorem 3.1], V'yugin [23, Theorem 4.1]; cf. Zvonkin and Levin [30, Theorem 3.3]). *The family of all r.e.b. semimeasures has the greatest semimeasure with respect to above defined quasiordering.*

Lemma 8.3.1. *A nonnegative function λ on Ξ is a semimeasure iff $(*)$ for every finite set $T \subset \Xi$: if $\forall p,\ q \in T\ \Omega_p \cap \Omega_q = \emptyset$ then $\sum_{t\in T} \lambda(t) \leqslant 1$.*

Proof. For semimeasures, the condition $(*)$ is evident. If, conversely, $(*)$ holds then we define the function μ on Ξ as follows. For $s \in \Xi$ let $R(s) = \{T \subseteq \Omega_s\colon \forall p,\ q \in T\ \Omega_p \cap \Omega_q = \emptyset\}$. Then $\mu(s) = \sup_{T\in R(s)} \sum_{t\in T} \lambda(t)$. The sought quasimeasure $\alpha(s)$ is defined by induction on the length of s: $\alpha(\Lambda) = 1$, $\alpha(s\hat{\ }0) = \mu(s\hat{\ }0)$, $\alpha(s\hat{\ }1) = \alpha(s) - \mu(s\hat{\ }0)$. The check of required properties is easy enough. □

Lemma 8.3.2. *Given a r.e.b. nonnegative function λ on Ξ, one can effectively construct a r.e.b. semimeasure λ' such that $\forall s \lambda'(s) \leqslant \lambda(s)$, and if λ is semimeasure itself then $\forall s \lambda'(s) = \lambda(s)$.*

Proof. For every finite step of the enumeration of the undergraph of λ, we have only finite subset of Ξ on which the current values of λ are positive. For that finite subset we can effectively check the condition $(*)$. If it does not hold then we stop enumeration of λ on the previous step. If $(*)$ holds we go on. □

Let $\mathbb{M}$ be the set of all r.e.b. nonnegative functions on Ξ. It is easy to see that $\mathbb{M}$ possesses a computable enumeration. That means that there exists a computable function $f: \mathbb{N} \to \mathbb{M}$ such that $\forall \lambda \in \mathbb{M}\ \exists n \in \mathbb{N}\ f(n) = \lambda$.

The function $\sum_{n=1}^{\infty} 2^{-n} f'(n)$ will do as the greatest r.e.b. semimeasure on Ξ (the operator $f(n) \mapsto f'(n)$ is defined in Lemma 8.3.2). □

Let us fix one of the greatest (in this sense) semimeasures and denote it by γ_A. Thus the recursively enumerable from below semimeasure γ_A possesses the quality that for every r.e.b. semimeasure γ there is a constant c satisfying the inequality $\gamma(s) \leqslant c\gamma_A(s)$ for all s.

Definition 8.4. The semimeasure γ_A is called the *a priori semimeasure*. The natural number $KA(s) = [-\mathrm{lb}(\gamma_A(s)) + 1]$ is called the *a priori entropy* of finite sequence s (where $[x]$ denotes the greatest integer less than or equal to x).

One may consider $KA(s)$ as a measure of complexity of a finite sequence s; one may interpret that measure, represented by natural number, as the amount of bits of information. The following question may arise:

will any results change if we consider another greatest semimeasure, say γ'_A (evidently there are infinitely many greatest semimeasures)?

By definition, the semimeasure γ'_A is *equivalent* to γ_A, that is there is a positive constant c such that $\gamma_A(s) \leqslant c\gamma'_A(s)$ and $\gamma'_A(s) \leqslant c\gamma_A(s)$ for all s. In a sense a priori semimeasure is defined to within a multiplicative constant and a priori entropy is defined to within an additive constant. In any case all presented statements are invariant in the sense mentioned above.

Now we are able to give the definition of chaoticness. Let a computable quasimeasure μ on Ξ be fixed. Let us define the *μ-complexity* $K = K_\mu$ of s by equality $K(s) = [-\mathrm{lb}(\mu(s)) + 1]$. It is clear that μ is r.e.b. semimeasure and therefore $KA(s) \leqslant K(s) + C$ for all s, where C is a constant not depending on s.

Definition 8.5. An infinite sequence $\mathbf{a}$ is called *μ-chaotic* if there is a constant C such that $KA(\mathbf{a}|n) \geqslant K_\mu(\mathbf{a}|n) - C$ for all n. Equivalently, $\gamma_A(\mathbf{a}|n) \leqslant c\mu(\mathbf{a}|n)$ for some c. (Here $\mathbf{a}|n$ stands for the tuple consisting of the first n terms of $\mathbf{a}$.) A sequence which is not μ-chaotic is called *μ-non-chaotic*.

Definition 8.6. Let us call η-chaotic sequences simply *chaotic*.

Recall that $\eta(s) = 2^{-\mathrm{lh}(s)}$ is the uniform measure. Thus, a sequence $\mathbf{a}$ is chaotic iff there is a constant C such that $KA(\mathbf{a}|n) \geqslant n - C$ for all n.

So for a μ-chaotic sequence the a priori entropy of its initial segments is equal to μ-complexity to within an additive constant.

Theorem 8.7 (majority theorem). *The set of μ-chaotic sequences has μ-measure* 1.

Proof. Let a positive ε be fixed. Let us cover the set Z of all μ-non-chaotic sequences with an open set G having μ-measure $\leqslant \varepsilon$. Let $C = 1 - \mathrm{lb}(\varepsilon)$. By definition, if $\mathbf{a}$ is a μ-non-chaotic sequence then there is $n = n(\mathbf{a})$ such that $KA(\mathbf{a}|n) \leqslant K_\mu(\mathbf{a}|n) - C$. This implies $\varepsilon\gamma_A(\mathbf{a}|n) \geqslant \mu(\mathbf{a}|n)$.

Let us denote by S the set of all $s \in \Xi$ such that $\varepsilon\gamma_A(s) \geqslant \mu(s)$ and no proper prefix s' of s satisfies the inequality $\varepsilon\gamma_A(s') \geqslant \mu(s')$. We have $Z \subseteq G = \bigcup_{s\in S} \Omega_s$ and on the other hand $\mu(G) = \sum_{s\in S} \mu(s) \leqslant \varepsilon \sum_{s\in S} \gamma_A(s) \leqslant \varepsilon$ because the semimeasure γ_A is not greater than some quasimeasure. $\square$

The following theorem [3] strengthens in a sense the preceding one. It states that the class of chaotic sequences is rich enough to generate all sequences. This theorem is not evident; for example, its analog for the notion of lawlessness is not true.

Theorem 8.8 (Gács). *Any infinite sequence is computable relative to some η-chaotic sequence.* (It is worthy to mention that there is a computable operator which gives *any* sequence as its output provided an *appropriate* chaotic sequence is taken as its input.)

We omit proof of Gács' Theorem.

The time is ripe to investigate connection between the chaoticness and the previous version of conception of randomness – the unpredictability.

Theorem 8.9. *All μ-chaotic sequences are μ-unpredictable.*

Proof. If G is an open subset of Ω and $\mu(G)<2^{-m}$ then the function $\lambda_{G,m}(s)=2^m\mu(\Omega_s \cap G)$ is a semimeasure. It suffices to confirm the condition $(*)$ from Lemma 8.3.1.

If G is enumerable then $\lambda_{G,m}$ is a r.e.b. semimeasure.

Lemma 8.9.1. *If $G(1), G(2), G(3), \ldots$ is a computable sequence of enumerable open subsets of Ω and $\forall n\, \mu(G(n))<2^{-n}$, then all sequences belonging to $\bigcap_{n=1}^{\infty} G(n)$ are μ-non-chaotic.*

Proof. If for some n $\mu(G(n))=0$, each sequence from $G(n)$ is μ-non-chaotic because the a priori semimeasure is positive on all arguments.

If $\forall n\, \mu(G(n))\neq 0$ then let us consider the function $\alpha=\sum_{n=1}^{\infty} 2^{-n}\lambda_{G(2n),2n}$. It is easy to see that α is a r.e.b. semimeasure. If a sequence $\mathbf{b}$ belongs to $\bigcap_{n=1}^{\infty} G(n)$ then $\forall n\, \exists k\, \alpha(\mathbf{b}|k)/\mu(\mathbf{b}|k)>2^n$ (k suits us if $\Omega_{\mathbf{b}|k} \subseteq G(2n)$). As $\gamma_A \geqslant c\alpha$ for some constant c, we obtain that $\mathbf{b}$ is μ-non-chaotic. □

To get Theorem 8.9 we will prove that any μ-predictable sequence $\mathbf{a}$ is not μ-chaotic. Let us fix a computable strategy σ giving the unbounded growth of Man's capital in the μ-game against $\mathbf{a}$. Due to Lemma 8.9.1, it suffices to construct (effectively) for every n such enumerable open G that $\mathbf{a}\in G$ and $\mu(G)<2^{-n}$. Let B be the set of such $\mathbf{b}\in\Omega$ that the strategy σ gets a capital larger than 2^n (at least once) in μ-game against $\mathbf{b}$. Evidently B is open and enumerable. We claim that $\mu(B)<2^{-n}$.

We will use now Definition 6.2.1.2. It is sufficient to prove that for any $S\subseteq \Xi$, if $\forall s,\ t\in S\ \Omega_s\cap\Omega_t=\emptyset$ and $\forall s\in S \mu(A_s)\neq 0$ then $\sum_{s\in S}\mu(A_s)V(s)\leqslant V_0$. It can be done analogously to the proof of Proposition 6.2.2. Note that unlike the present case, in 6.2.2 there was proved the equality. The cause is that an additional condition $\bigcup_{s\in S}\Omega_s=\Omega$ was present in Proposition 6.2.2 but is absent now. □

Corollaries 8.10. 1. *Theorem* 7.2 *and hence its Corollary* 7.3 (*i.e. the majority theorem for unpredictability*) *can be obtained as some relativized versions of Theorems* 8.9 *and* 8.7. *To get a proof of Theorem* 7.2 *we have to consider the strategy σ from that theorem as an* oracle *for the notions connected with computability, i.e. for chaoticness and unpredictability. It means that one needs to relativize Theorems* 8.9 *and* 8.7 *to that oracle.*

2. *A further consequence is a new proof of Corollary* 3.5 (*i.e. majority theorem for η-stochastic sequences*).

3. *An immediate consequence of Theorem* 8.9 *and Theorem* 7.5 *is the theorem of distinguishing for chaoticness*: no sequence is chaotic with respect to two inconsistent computable measures.

4. *Using Corollary* 7.7 *we obtain that* there are stochastic non-chaotic sequences.

Open question 8.11. Are the notions of chaoticness and unpredictability equivalent?

9. Theorems on complexity deficiency and other features of chaotic, unpredictable and stochastic sequences

Let us consider now the following question. As the class of non-chaotic sequences contains a stochastic sequence (Corollary 4 of Theorem 8.9) and it may well be true that class contains an unpredictable sequence (this is the open problem posed in the end of preceded section), it is natural to ask: for what functions d there is a stochastic (or unpredictable) sequence $\mathbf{a}$ satisfying the inequality $KA(\mathbf{a}|n)<n-d(n)$ for any sufficiently large n. Such function d is often called *complexity deficiency*. In a more narrow sense the term *complexity deficiency* denotes the difference $n-KA(\mathbf{a}|n)$. Let us remind that if it is not stated otherwise we assume measure to be uniform. It is the uniform measure η that is considered in this section.

We regard the two first theorems of this section, which connect the properties of unpredictability and stochasticness with the behavior of the complexity deficiency, as hardly not the most important ones in the present paper. The second of those theorems yields the solution of a famous problem of recursion theory which is related with the correction of one false assertion posed in the 1960s by A.N. Kolmogorov (see remark after the formulation of Theorem 9.2).

Theorem 9.1 (An.A. Muchnik). *Let D be a computable total function. An unpredictable (i.e. η-unpredictable) sequence $\mathbf{a}$ satisfying, for any sufficiently large n, the inequality $KA(\mathbf{a}|n)\leqslant n-D(n)$ exists if and only if the function D is bounded.*

Proof. *"IF" part*: Let us prove first that if D is bounded then there is an unpredictable sequence $\mathbf{a}$ satisfying the inequality $KA(\mathbf{a}|n)<n-D(n)$ for any sufficiently large n. Let a constant C be given. We have to construct an unpredictable sequence $\mathbf{a}$ such that $KA(\mathbf{a}|n)<n-C$ for any sufficiently large n.

Let us denote $Z_l=\langle 0,\dots,0\rangle$ (l zeros) and $Y_l=Z_l\hat{\ }1$. Let us define semimeasure μ on Ξ by following conditions:

$\mu(Y_{2l}\hat{\ }t)=2^{-l-\mathrm{lh}(t)-1}$ for every $t\in\Xi$; if a tuple s has not the form $Y_{2l}\hat{\ }t$ then $\mu(s)=0$.

Of course μ is recursively enumerable from below (r.e.b.), therefore $\exists P\ \forall s\in\Xi\ KA(s) <-\mathrm{lb}(\mu(s))+P$.

For every l, there exist unpredictable sequences having the prefix Y_{2l} (because almost all sequences are unpredictable). For each such sequence $\mathbf{b}$ and for every $n>2l$ we have $KA(\mathbf{b}|n)<n-l+P$. Now we take $l>C+P$. $\square$

"ONLY IF" part: Let us prove now that if a function D is unbounded (and computable) and a sequence $\mathbf{a}$ satisfies the inequality $KA(\mathbf{a}|n)<n-D(n)$ for any sufficiently

large n then $\mathbf{a}$ is predictable. We will construct two computable strategies such that one of them will be winning in the game against the sequence $\mathbf{a}$.

Let us partition the natural row $\mathbb{N}$ into the semiintervals $N_k=[n_k,n_{k+1})$ called *zones*. The sequence $n_1,n_2,n_3,\ldots$ will be increasing and computable. Numbers n_k are defined by induction on k. We set $n_1=1$. Assume that n_k is already defined. Let us define n_{k+1}. As D is unbounded, there is n such that $n>n_k$ and $D(n)>3n_k+k$. Let us take the least such n and set $n_{k+1}=n+1$.

It is very important that in the definition of n_k we have not used the sequence $\mathbf{a}$ itself, but only the function D and its computability.

We denote by $\mathbf{a}|N_k$ the finite sequence $\langle\mathbf{a}(n_k),\mathbf{a}(n_k+1),\ldots,\mathbf{a}(n_{k+1}-1)\rangle$.

Lemma 9.1.1. *For some constant C*

$$\forall s,\ t\in\Xi\, KA(t)<KA(s\hat{}t)+2\,\mathrm{lh}(s)+C.$$

Proof. We can easily see that the function $\mu(t)=\sum_{s\in\Xi}2^{-2\mathrm{lh}(s)-1}\gamma_A(s\hat{}t)$ is a r.e.b. semimeasure. Therefore $\exists c\,\forall t\in\Xi\ \mu(t)<c\gamma_A(t)$. Taking the logarithm of the previous inequality, we have

$$\forall s,\ t\in\Xi KA(t)<KA(s\hat{}t)+2\,\mathrm{lh}(s)+\mathrm{lb}(c)+3. \qquad\square$$

Due to Lemma 9.1.1 we have

Proposition 9.1.2. $KA(\mathbf{a}|N_k)<n_{k+1}-n_k-k+C.$

The idea of construction of two computable strategies is as follows. As the a priori semimeasure is recursively enumerable from below, the complexity function $KA(s)$ is recursively enumerable from above. This means that there is a (total) computable function $\theta:\mathbb{N}\to\mathbb{N}\times\Xi$ such that

$$KA(s)=\inf_{j\in\mathbb{N}}\{\theta_1(j):\ \theta_2(j)=s\},\ \text{where}\ \theta(j)=\langle\theta_1(j),\theta_2(j)\rangle.$$

We denote by $J_k(\mathbf{a})$ the least j such that $\theta_2(j)=\mathbf{a}|N_k$ and $\theta_1(j)<n_{k+1}-n_k-k+C$.

Now we will give the most important definition. Let us say that $\mathbf{a}$ is a sequence of *even type* if for infinitely many k the inequality $J_{2k-1}(\mathbf{a})\leqslant J_{2k}(\mathbf{a})$ holds; let us say that $\mathbf{a}$ is a sequence of *odd type* if for infinitely many k the inequality $J_{2k-1}(\mathbf{a})\geqslant J_{2k}(\mathbf{a})$ holds. Clearly, $\mathbf{a}$ has at least one of these two types. We will now define two computable Man's strategies – *even* and *odd* ones – such that if $\mathbf{a}$ has even type then the even strategy is winning, and if $\mathbf{a}$ has odd type then the odd strategy is winning (in the game against $\mathbf{a}$).

9.1.3. *The even strategy*. Man plays in zones N_k, $k\in\mathbb{N}$, in the following order: $N_2,N_1,N_4,N_3,N_6,N_5,\ldots$. The expression "game in zone N_k" means that Man predicts the values of terms of the sequence $\mathbf{a}$ with numbers from N_k in natural order: n_k,n_k+1 and so on; thus, the total number of moves in zone N_k is $l_k=n_{k+1}-n_k$.

Let us describe Man's action in zones N_{2k} and N_{2k-1}.

Zone N_{2k}. Man makes l_{2k} moves, betting on ith move ($n_{2k} \leqslant i < n_{2k+1}$) zero on an arbitrary value of the ith term. That means that Man finds out $\mathbf{a}|N_{2k}$ without changing his capital. This knowledge allows Man to compute $J_{2k}(\mathbf{a})$.

Zone N_{2k-1}. At first Man constructs a subset $S(2k-1)$ of the set $I(2k-1)$ of all sequences of length l_{2k-1} defined as follows:

$$S(2k-1) = \{s \in I(2k-1)\colon \exists j[j \leqslant J_{2k}(\mathbf{a})\,\&\,\theta_2(j)=s \\ \&\ \theta_1(j) < l_{2k-1} - 2k + 1 + C]\}.$$

That is, $S(2k-1)$ is the set of all $s \in I(2k-1)$ such that the values $\theta(j)$, $j \leqslant J_{2k}(\mathbf{a})$, allow to conclude that $KA(s) < l_{2k-1} - 2k + 1 + C$.

Each $s \in S(2k-1)$ has complexity $KA(s) \leqslant l_{2k-1} - 2k + C$ and therefore $\gamma_A(s) \geqslant \exp_2[-(l_{2k-1} - 2k + C)]$. As the semimeasure γ_A is majorized by a quasimeasure, the set $S(2k-1)$ has at most $\exp_2(l_{2k-1} - 2k + C)$ elements. This means that the uniform measure of the set $S(2k-1)$ is at most 2^{-2k+C}.

On the other hand, if $J_{2k-1}(\mathbf{a}) \leqslant J_{2k}(\mathbf{a})$ then $\mathbf{a}|N_{2k-1} \in S(2k-1)$, therefore Man has much information about $\mathbf{a}|N_{2k-1}$ (because $S(2k-1)$ is "small").

Having $S(2k-1)$ Man begins to play in zone N_{2k-1}. He does it as it has been described in the proof of item 1 of Theorem 6.1.1, where the measure (φ) is uniform, the number of moves (m) is equal to l_{2k-1}, the set (S) is equal to $S(2k-1)$, and initial capital is equal to the half of capital which Man had before the beginning of the game on $(2k-1)$th zone.

If $\mathbf{a}|N_{2k-1} \notin S(2k-1)$ then Man's capital decreases at most twice during the game on zone N_{2k-1}.

If $\mathbf{a}|N_{2k-1} \in S(2k-1)$ then Man's capital increases at least 2^{2k-C-1} times.

Recall that after the game on zones N_{2k}, N_{2k-1} Man turns to zones N_{2k+2}, N_{2k+1}.

The definition of even strategy is completed. It is clear that the described strategy is computable.

Lemma 9.1.4 *If* $\mathbf{a}$ *is a sequence of even type then in the game against* $\mathbf{a}$ *Man wins via the even strategy.*

Proof. At the end of the game on the first k pairs of zones the initial capital decreases at most 2^k times.

If $J_{2k-1}(\mathbf{a}) \leqslant J_{2k}(\mathbf{a})$ then the current capital increases during the game on the kth pair of zones at least 2^{2k-C-1} times. So the initial capital increases in this case at least 2^{k-C} times. The condition of Lemma tells that the set of such cases is infinite. □

9.1.3.* *The odd strategy.* The zones are disposed in the order $N_1, N_2, N_3, N_4, N_5, N_6, \ldots$, i.e. with inversion in every pair compared to the even strategy. Playing in zone N_{2k-1} Man makes zero bets, he keeps his capital and finds out $\mathbf{a}|N_{2k-1}$. This enables him to

compute $J_{2k-1}(\mathbf{a})$. Then, as in above case, the sets $I(2k)$ and $S(2k)$ are defined and $|S(2k)|<2^{-2k+C}|I(2k)|$. The strategy of playing on zone N_{2k} is defined in similar way; this strategy increases the current capital at least 2^{2k-C-1} times if $J_{2k-1}(\mathbf{a})\geqslant J_{2k}(\mathbf{a})$ and decreases the current capital at most twice otherwise.
Finally we can prove

Lemma 9.1.4.* *If* $\mathbf{a}$ *is a sequence of odd type then in the game against* $\mathbf{a}$ *Man wins via the odd strategy.*

Theorem 9.2 (An.A. Muchnik). *Let D be a computable total function. Then the following are equivalent*:

(1) *there exists a stochastic* (*i.e. η-stochastic*) *sequence* $\mathbf{a}$ *satisfying the inequality* $KA(\mathbf{a}|n)\leqslant n-D(n)$ *for any sufficiently large* n;

(2) *for every* $c>0$ *the inequality* $D(n)<cn$ *holds for any sufficiently large* n.

Remark. The assertion posed by Kolmogorov which was mentioned in the beginning of this Section 9 is the following: there exists a stochastic sequence x such that $H(x^l)=\mathrm{O}(\log l)$. That assertion, without a proof, can be found in the last paragraphs of article 2 of [6, 7]. We observe here Kolmogorov's notations yet are going to explain them: x^l is the same as $x|l$ (in our notations), and H stands for the *simple entropy*. (On various kinds of entropy see, e.g., [19], and also [22].) The simple entropy differs from a priori entropy but not too much: $|H(s)-KA(s)|=\mathrm{O}(\log|s|)$ for $s\in\Xi$. Therefore, Kolmogorov's statement asserts, in fact, the existence of a stochastic sequence $\mathbf{a}$ such that $KA(\mathbf{a}|n)=\mathrm{O}(\log n)$. However, this is refuted by Theorem 9.2. The situation with Kolmogorov's assertion was discussed in [20, no. 2.6.5] and in Section 1.6 of [8, 9]. A slightly weakened version of Theorem 9.2 was published in [21, no. 6.2.3]:

Theorem 9.2′ (An.A. Muchnik). *Let ω be an infinite sequence of zeros and ones and suppose that the entropy of an initial segment of ω having length n does not exceed αn for some $\alpha<1$ and for any sufficiently large n. Then ω is not stochastic.*

Proof of Theorem 9.2. Suppose first that for every $c>0$, $D(n)<cn$ for any sufficiently large n. We have to construct a stochastic sequence $\mathbf{a}$ satisfying the inequality $KA(\mathbf{a}|n)\leqslant n-D(n)$ for any sufficiently large n. It is clear that without loss of generality we may assume that D tends to infinity (if this is not the case then we can add $\mathrm{lb}(n)$).

Let us define d_n such that $d_n^2=8\max_{n\leqslant k\leqslant 2n}D(k)/k$. Obviously, the sequence d_n is computable and $d_n\to 0$ (as for any $\varepsilon>0$ there is such k_ε that $D(k)\leqslant\varepsilon k$ for all $k\geqslant k_\varepsilon$). We claim that $\sum_{k=1}^n d_k^2\geqslant 4D(n)$ for all n. Indeed, the definition of d_n yields that for all m, if $n/2\leqslant m\leqslant n$ then $d_m^2\geqslant 8D(n)/n$; therefore $\sum_{k=1}^n d_k^2\geqslant(n/2)[8D(n)/n]$. Thus we have $\sum_{k=1}^\infty d_k^2=\infty$.

Let us consider the quasiuniform measure μ defined by the sequence d_n (see the statement of Theorem 5.1). The μ-measure of the set of all η-stochastic sequences is 1 (according to 5.4 and 3.14). On the other hand, μ-measure of the set of all μ-chaotic sequences is also equal to 1 (Theorem 8.7). Therefore, there are sequences which are μ-chaotic and η-stochastic simultaneously. Let us fix such a sequence $\mathbf{a}$ and prove that $KA(\mathbf{a}|n) \leqslant n - D(n)$ for any sufficiently large n.

To this end we consider the quasiuniform measure μ' defined by the sequence $d_n/2$. Then $\mu'(\mathbf{a}|n)^2 \geqslant \mu(\mathbf{a}|n)2^{-n}\prod_{k=1}^{n}(1+d_k^2/2)$ (see the proof of 5.1). As γ_A is a maximal semimeasure, there is a constant $c'>0$ such that $\gamma_A(\mathbf{a}|n) \geqslant c'\mu'(\mathbf{a}|n)$ for all n. As $\mathbf{a}$ is μ-chaotic, there is another constant $c''>0$ such that $\mu(\mathbf{a}|n) \geqslant c''\gamma_A(\mathbf{a}|n)$ for all n. Therefore,

$$\gamma_A(\mathbf{a}|n) \geqslant c2^{-n}\prod_{k=1}^{n}(1+d_k^2/2), \quad \text{where } c=c''(c')^2;$$

$$KA(\mathbf{a}|n) \leqslant n - \sum_{k=1}^{n} \mathrm{lb}(1+d_k^2/2) - C, \quad \text{where } C=\mathrm{lb}(c).$$

As $2<e$, then $\mathrm{lb}(1+\alpha)>\alpha$ for small positive α. Consequently,

$$KA(\mathbf{a}|n) \leqslant n - \sum_{k=1}^{n} d_k^2/2 - C - C' \leqslant n - 2D(n) - C - C',$$

where the term C' arises because a finite number of d_k's may not satisfy the inequality $\mathrm{lb}(1+(d_k^2/2)) \geqslant d_k^2/2$. Thus, $KA(\mathbf{a}|n) \leqslant n - D(n)$ for any sufficiently large n (because D tends to infinity).

Let us prove now the inverse implication. Assume that $\mathbf{a}$ is a stochastic sequence satisfying the inequality $KA(\mathbf{a}|n) \leqslant n - D(n)$ for any sufficiently large n and assume that D is computable. We have to prove that D grows more slowly than any increasing linear function.

Assume the contrary: there is $\alpha>0$ such that $D(n)>\alpha n$ for infinitely many n. Without loss of generality we may assume that α has the form $\alpha=M^{-1}$, where $M\in\mathbb{N}$.

Let us define the computable increasing sequence n_k of natural numbers in any way to satisfy the following inequalities:

(1) $D(n_{k+1}-1)>\alpha(n_{k+1}-1)$.

(2) $n_{k+1}>2M(3n_k+C)$ where the constant C is taken from Lemma 9.1.1.

(3) $n_{k+1}>128M^2n_k$.

Let us denote by N_k the semi-interval $[n_k, n_{k+1})$. From Lemma 9.1.1 and from the conditions (1), (2) it follows that

$$\begin{aligned} KA(\mathbf{a}|N_k) &< KA(\mathbf{a}|(n_{k+1}-1)) + 2n_k + C < n_{k+1}(1-\alpha) + 2n_k + C \\ &< (n_{k+1}-n_k)(1-\alpha/2). \end{aligned}$$

Further, we define the value $J_k(\mathbf{a})$ and the notions of a sequence of even and odd type in the way similar to the way used in the proof of Theorem 9.1. By definition $J_k(\mathbf{a})$ is the least j satisfying conditions: $\theta_2(j)=\mathbf{a}|N_k$ and $\theta_1(j)<l_k(1-\alpha/2)$ where $l_k=n_{k+1}-n_k$.

However, now we need a finer classification of sequences and rules than the classification from the proof of Theorem 9.1. We will have a second parameter – a pair of integers $z=\langle z_1,z_2\rangle$ from the set $\{0,1\}\times[1,4M]$. Let us begin with the rules of choice of subsequences.

9.2.1. *Even* (z) *rule*. As to level of zones (about the term "zone" see the proof of Theorem 9.1) the rule acts similar to the even strategy from the proof of Theorem 9.1. In particular, any even rule scans zones in the order $N_2,N_1,N_4,N_3,N_6,N_5,\ldots$, and every "even" zone N_{2k} is scanned only to find out $\mathbf{a}|N_{2k}$ and no value a_i, $i\in N_{2k}$, is chosen to be included in the subsequence.

Then Man computes the set $S=S(2k-1)\subseteq I(2k-1)$ (where $I(2k-1)$ is the set of all binary sequences of length $l_{2k-1}=n_{2k}-n_{2k-1}$ and $S(2k-1)$ is defined in the proof of Theorem 9.1). Obviously, $S(2k-1)$ has at most $\exp_2[l_{2k-1}(1-\alpha/2)]$ elements and $\mathbf{a}|N_{2k-1}\in S(2k-1)$ if $J_{2k-1}(\mathbf{a})\leqslant J_{2k}(\mathbf{a})$ (in this case we say that $(2k-1)$th zone has *even type*).

Let us consider the strategy σ constructed in the proof of item 1 of Theorem 6.1.3 and let us fix the values of parameters of that theorem: $m=l_{2k-1}$, $S=S(2k-1)$. The winnings of σ against each element of S are not less than $\alpha l_{2k-1}/2$. Let us define new strategy σ' which bets on the same values as σ, and if β is bet of σ then $[4M\beta]/4M$ is bet of σ'. Clearly the winnings of σ' may differ from the winnings of σ not more than by $l_{2k-1}/4M=\alpha l_{2k-1}/4$. So the winnings of σ' are not less than $\alpha l_{2k-1}/4$.

Now let us divide the strategy σ' into $8M$ substrategies σ_z, $z\in\{0,1\}\times[1,4M]$, as follows. If σ' bets on z_1 and if its bet, multiplied by $4M$, is equal to z_2, then σ_z does the same. In other cases σ_z bets zero. For every sequence t from $S(2k-1)$ there is such z that the winnings of substrategy σ_z against t are not less than $(\alpha l_{2k-1}/4)/8M=\alpha^2 l_{2k-1}/32$. We will call such z for $\mathbf{a}|N_{2k-1}$ *the even type of* $(2k-1)$*th zone.*

The even (z) rule takes a term in subsequence iff the bet of σ_z is positive. As the values of all positive bets of σ_z are equal to $z_2/4M\leqslant 1$, then the difference between the number of (z_1) and the number of $(1-z_1)$ in chosen from $\mathbf{a}|N_{2k-1}$ subsequence is not less than the winnings of σ_z. If $(2k-1)$th zone has the even type z, then the even (z) rule guarantees, that the number of (z_1) chozen from $\mathbf{a}|N_{2k-1}$ is greater than

$$\begin{aligned} l_{2k-1}/2+\alpha^2 l_{2k-1}/64 &= n_{2k}(1/2+\alpha^2/64)-n_{2k-1}(1/2+\alpha^2/64)\\ &> n_{2k}(1/2+\alpha^2/128)+n_{2k}\alpha^2/128-n_{2k-1}\\ &> n_{2k}(1/2+\alpha^2/128) \end{aligned}$$

(due to condition (3)). The length of subsequence chosen from $\mathbf{a}|n_{2k}$ is not greater than n_{2k}. Thus, we see that the law of large numbers is not true for the chosen subsequence.

The set of all types and rules is finite. Therefore, it is sufficient to use such rule that infinitely many zones have corresponding type for sequence $\mathbf{a}$.

9.2.1.* *Odd* (z) *rule* scans the zones in the order $N_1,N_2,N_3,N_4,N_5,N_6,\ldots$. The difference with the even (z) rule is that in every pair N_{2k-1},N_{2k} the roles of zones are changed, as in the proof of Theorem 9.1. The notion of odd type of $(2k)$th zone is introduced in the similar way as for even type. The details are left to the reader.

The computability of all used strategies and rules is obvious. □

Theorem 9.3 (An.A. Muchnik). *For every computable strategy there exist a number $c>0$ and a sequence* $\mathbf{a}$ *such that*

(1) *for any measure μ the strategy does not win in the μ-game against* $\mathbf{a}$; *and*

(2) $KA(\mathbf{a}|n) \leqslant c\,\mathrm{lb}(n)$ *for all $n>1$.*

If the given strategy is monotone or if for every number i, at some moment of the game, it predicts the ith term then there exists a computable sequence $\mathbf{a}$ *satisfying* (1) *and* (2).

This means that there is no universal computable strategy (i.e. strategy which wins against all predictable sequences), because the sequence $\mathbf{a}$ is predictable due to Theorem 9.1. However, for every countable set of strategies there exists universal continuous strategy, if we do not require this strategy to belong to that set.

Proof. Let a computable strategy σ be given. Let us define by induction on k the sequence of indices n_k and the sequence of terms a_{n_k} as follows:

(1) If $k=1$ then n_1 is the number of the first term predicted by strategy σ and $a_{n_1} = 1 - \tilde{a}_{n_1}$, where $\tilde{a}_{n_1}$ is the predicted value.

(2) If $k \geqslant 2$ then n_k is the number of kth term predicted by strategy σ in the case when the true values of $(k-1)$ before predicted terms are $a_{n_1}, a_{n_2}, \ldots, a_{n_{(k-1)}}$ and $a_{n_k} = 1 - \tilde{a}_{n_k}$, where $\tilde{a}_{n_k}$ is the predicted value.

Of course, the mapping $n \mapsto a_n$ is computable and possibly is not totally defined. Let $a_n = 0$ for all n that are not included in the sequence n_k.

It is clear that strategy σ does not win in the game against the defined sequence $\mathbf{a}$ because all predictions are false. To complete the proof it remains to verify the inequality $KA(\mathbf{a}|n) \leqslant 4\mathrm{lb}(n) + C$ for all n (thus, we can take $c = 4 + C$). It is sufficient to construct a r.e.b. semimeasure ν such that $\nu(\mathbf{a}|n) \geqslant n^{-4}$ for all $n \geqslant n_1$.

Let us define an auxillary function $\nu: \Xi \to \mathbb{Q}$. Let $s \in \Xi$ be a sequence of length n. Assume that among the numbers $n' \leqslant n$ there are exactly m numbers of the form n_k. Let these numbers be $n_{k(1)}, n_{k(2)}, \ldots, n_{k(m)}$, where $k(1) < k(2) < \cdots < k(m)$. We set $\nu(s) = \nu_1(s) + \nu_2(s) + \cdots + \nu_m(s)$, where $\nu_l(s) = n^{-4}$ if

(1) $s_{n_{k(i)}} = a_{n_{k(i)}}$ for all $i \leqslant l$ and

(2) $s_j = 0$ for all $j \leqslant n$ such that $j \notin \{n_{k(i)}\colon i \leqslant l\}$

and $\nu_l(s) = 0$ else.

It is clear that ν is recursively enumerable from below and that $\nu(\mathbf{a}|n) \geqslant n^{-4}$ if $n \geqslant n_1$. To prove that ν is a semimeasure we have to construct a majorizing quasimeasure. Let us note that $\nu(s) \neq 0$ only for those sequences s of length n that satisfy (1) and (2) for some $l \leqslant m$. The number of such sequences is at most n (as $m \leqslant n$) and for any of them $\nu(s) \leqslant nn^{-4} = n^{-3}$. Consequently, the sum of values $\nu(s)$ over all binary s of length n is at most n^{-2}. Hence $\sum_{s \in \Xi} \nu(s)$ is finite.

Let us define now $\tau(s) = \sum_t v(t)$, where the sum is taken over all extensions t of the sequence s (including s itself). The function τ is nonnegative and satisfies the inequality $\tau(s) \geqslant \tau(s\hat{\ }0) + \tau(s\hat{\ }1)$ for all $s \in \Xi$. To convert this inequality into equality let the value $\theta(s)$ be defined by induction on $\mathrm{lh}(s)$ as follows:

(a) $\theta(\Lambda) = \tau(\Lambda)$ (where Λ is the empty sequence),

(b) $\theta(s\hat{\ }1) = \tau(s\hat{\ }1)$ and $\theta(s\hat{\ }0) = \theta(s) - \theta(s\hat{\ }1)$.

We can easily verify that θ is really a quasimeasure majorizing τ and, therefore, majorizing v. Thus, the proof of the first assertion of theorem is completed.

The second assertion is in fact also proved: if the given strategy is monotone then the assertion is evident and if the given strategy in the game against any sequence predicts all its terms, then the sequence constructed in the proof is computable. □

The next theorem yields a computable strategy universal for the class of all computable sequences for the game with uniform measure η (note that due to Theorem 9.3 this strategy cannot be monotone).

Theorem 9.4 (An.A. Muchnik). *There is a computable strategy which wins in the η-game against any computable sequence.*

Proof. Let us fix a partial computable 0–1-valued function $f_k(n)$ of two arguments universal in the following sense: for every partial computable function $f : \mathbb{N} \to \{0,1\}$ there are infinitely many k such that $f(n) = f_k(n)$ for all n (as it is usual in the study of partial functions, the equality means that if one hand side is defined then the other hand side is also defined and both hand sides are equal). The set $F = \{\langle k,n,i\rangle\colon f_k(n) = i\}$ is recursively enumerable; there is a tuple of total computable functions $\tilde{k}, \tilde{n}, \tilde{i}$ such that $F = \{\tau(l) = \langle \tilde{k}(l), \tilde{n}(l), \tilde{i}(l)\rangle\colon l \in \mathbb{N}\}$

We define by induction on k the sequence of numbers n_k: $n_1 = 1$ and $n_{k+1} = n_k + k$ for all k. Let N_k denote the semi-interval $[n_k, n_{k+1})$ in $\mathbb{N}$.

We define $l_k = \min\{l\colon \forall n \in N_k\ \exists l' \leqslant l\ [k = \tilde{k}(l'),\ n = \tilde{n}(l'),\ f_k(n) = \tilde{i}(l')]\}$. We can say that l_k is the number of operations which is spent for computing of $f_k(n)$ for all $n \in N_k$. (As f_k may be non-total, l_k can be equal to $+\infty$.) We define $K = \{k \in \mathbb{N} \mid l_k < +\infty\}$. Let $k[1], k[2], k[3], \ldots$ be the recursive enumeration of K without repetitions. The mapping $m \mapsto k[m]$ is computable and is total (because there are infinitely many k such that f_k is total).

Now let us turn to the description of the strategy. Let the initial capital be 1. The game is divided into infinitely many steps: step 1, step 2, step 3,.... The step m is as follows. Man computes $k = k[m]$, l_k, all tuples $\tau(l) = \langle \tilde{k}(l), \tilde{n}(l), \tilde{i}(l)\rangle$, $l \leqslant l_k$, and using them Man computes $f_k(n)$ for all $n \in N_k$. Man takes 2^{-k} as initial capital for this step. Then he makes k moves, on nth move he bets his current capital on the value $f_k(n_k + n - 1)$ of the term with number $(n_k + n - 1)$.

Obviously, if $a_n = f_k(n)$ for all $n \in N_k$, then Man increases 2^k times the capital 2^{-k} on the step m. In this case the current capital after executing the step number m is increased by $1 - 2^{-k} \geqslant 0.5$. In other case, i.e. if $a_n \neq f_k(n)$ for some $n \in N_k$, Man loses the capital 2^{-k}. Therefore the sum of losses during all steps is not greater than 1. On the other hand, if the sequence $\mathbf{a}$ is computable, then for infinitely many k the equality $a_n = f_k(n)$ holds for all n, hence on infinite number of steps Man adds a capital greater than $\frac{1}{2}$. This means that the capital is unbounded and even tends to infinity in the game against $\mathbf{a}$. $\square$

Theorem 9.5 (An.A. Muchnik). *For every computable monotone function f tending to infinity there exists a sequence $\mathbf{a}$ which Man cannot win against if he uses only monotone strategies (i.e. $\mathbf{a}$ is monotonically unpredictable) and such that $KA(\mathbf{a}|n) < f(n)$ $\mathrm{lb}(n)$ for any sufficiently large n.*

Proof. We shall mean that monotone strategies bet zero on omitted terms. Let us consider an auxillary computable monotone function $g : \mathbb{N} \to \mathbb{N}$ tending to infinity such that $g(n) < \mathrm{lb}(n+1)$. Its values will be fixed at the end of the proof. For any $i \in \mathbb{N}$ let $h(i)$ be equal to the least n such that $g(n) > i$.

Now we will define three sequences by simultaneous induction on n. These sequences are: $s_n \in I^n$, $M_n \subset [1, g(n)]$, $v_n \in \mathbb{Q}$. For $n=0$ we set: $s_0 = \Lambda$, $M_0 = \emptyset$, $v_0 = 1$. Assume that we have already defined nth terms. For $(n+1)$ we have the following. Let us consider values of first $g(n)$ monotone strategies (in some natural numeration) on the argument s_n. M_{n+1} is the union of M_n and the set of numbers from $[1, g(n)]$ of those monotone strategies which are not defined on s_n. Let $V_k^i(t)$ be the current capital of ith monotone strategy after kth move if $\mathbf{a}|k = t$. Let us consider the next two sums:

$$u_0 = \sum_i 2^{-h(i)-i} V_{n+1}^i(s_n \hat{\ } 0) \quad \text{and} \quad u_1 = \sum_i 2^{-h(i)-i} V_{n+1}^i(s_n \hat{\ } 1),$$

where $i \in [1, g(n)] \backslash M_{n+1}$. If $u_0 \leqslant u_1$ then $v_{n+1} = u_0$, $s_{n+1} = s_n \hat{\ } 0$; and if $u_1 < u_0$ then $v_{n+1} = u_1$, $s_{n+1} = s_n \hat{\ } 1$.

It is easy to check the next inequalities

$$v_{n+1} \leqslant v_n \qquad \text{for } g(n) = g(n-1) \text{ and}$$

$$v_{n+1} \leqslant v_n + 2^{-g(n)} \quad \text{for else.}$$

For this check we use the fact that $(V_{k+1}^i(t \hat{\ } 0) + V_{k+1}^i(t \hat{\ } 1))/2 = V_k^i(t)$ (for all $i \in \mathbb{N}$, $t \in \Xi$ and $k = |t|$), and that $V_k^i(t) \leqslant 2^k$. So $v_n \leqslant v_0 + \sum_{i=1}^{\infty} 2^{-i} = 2$ for all n, therefore $V_n^i(s_n) \leqslant 2 \cdot 2^{h(i)+i}$ for all $i \notin M_n$ and $n \geqslant h(i)$.

We define $\mathbf{a}|n = s_n$. If ith monotone strategy is defined in the game against $\mathbf{a}$ on all moves, then $V_n^i(\mathbf{a}|n)$ is bounded. If ith monotone strategy is not defined in the game against $\mathbf{a}$ on some move, then i will belong to appropriate M_k. In all cases ith monotone strategy does not win against sequence $\mathbf{a}$.

It remains to bound the entropy of $\mathbf{a}$. Let us define a r.e.b. semimeasure μ as follows. First, we define by induction on $|t|$ an auxillary r.e.b. function $\mu_L(t)$ with two arguments: $t \in \Xi$ and L which is a tuple of length $|t|$ and kth element of L is a subset of $[1, g(k)]$ (where $L(k) \subset L(k+1)$). We set $\mu_L(\Lambda) = 1$. Assume that for all L values $\mu_L(t)$ are defined. Let n be equal to $|t|$ and P be a tuple of length $(n+1)$. We want to define $\mu_P(t \hat{\ } 0)$ and $\mu_P(t \hat{\ } 1)$. At first we compute $(n+1)$th moves of monotone strategies with numbers from $[1, g(n)]\backslash P(n)$ in the game against t. If one of such strategies is not defined on $(n+1)$th move in the game against t then $\mu_P(t\hat{\ }0) = \mu_P(t\hat{\ }1) = 0$. In another case we consider two sums: $u_0 = \sum_i 2^{-h(i)-i} V^i_{n+1}(t\hat{\ }0)$ and $u_1 = \sum_i 2^{-h(i)-i} V^i_{n+1}(t\hat{\ }1)$ where $i \in [1, g(n)]\backslash P(n)$. Let u_j be the least of two sums. Now we set $\mu_P(t\hat{\ }(1-j)) = 0$ and $\mu_P(t\hat{\ }j) = \mu_M(t) \cdot \alpha$ where $M = P|n, \alpha = 1 - (n+1)^{-2}$ if $P(n+1) = P(n)$ and $\alpha = (n+1)^{-3}$ if $P(n+1) \neq P(n)$.

It is easy to verify that $\mu_L(t)$ is really a r.e.b. function and that $\mu(t) = \sum_M \mu_M(t)$ is a semimeasure (use that $g(n) < \mathrm{lb}(n+1)$ and induction on n). For the sequence M_n, defined in the construction of **a**, we have

$$\mu(\mathbf{a}|k) \geqslant \mu_P(\mathbf{a}|k) \geqslant (k^{-3})^{g(k)} \prod_{m=2}^{\infty} (1 - m^{-2}),$$

where $P(n) = M_n$ for all n. Therefore $KA(\mathbf{a}|K) \leqslant 3g(k)\mathrm{lb}(k) + C$ where C does not depend on k. Hence, $KA(\mathbf{a}|k) \leqslant f(k)\mathrm{lb}(k)$ for any sufficiently large k, if $g(k) < f(k)/3$. □

Corollary 9.5.1. Theorems 9.2 and 9.5 (for $f(n) = \mathrm{lb}(n)$) *imply that there exists a monotonically unpredictable sequence which is not stochastic.*

By Theorem 7.4 the sequence constructed in 9.5.1 is also predictable.

Note that there exist a monotonically predictable sequence with a large a priori entropy (Corollary 7.7, Theorem 9.2) and a montonically unpredictable sequence with a small a priori entropy (Theorem 9.5). The fact that monotonical predictability is not correlated with a priori entropy is rather interesting.

It is interesting that some important results (e.g. Theorem 9.7) about randomness with respect to a measure μ can be obtained for sequences which are chaotic with respect to some (another) measure ν.

Definition 9.6. Let us call a sequence **a** *natural* if there is a computable measure ν such that **a** is ν-chaotic.

Theorem 9.7 (An.A. Muchnik). *For any computable measure μ, the notions of μ-chaoticness and μ-unpredictability are equivalent in the class of natural sequences: a natural sequence is μ-unpredictable if and only if it is μ-chaotic.*

Proof. For "*if*" see Theorem 8.9.

"*Only if*". Let us suppose that a sequence **a** is a natural and let ν be a computable measure such that **a** is ν-chaotic. Let **a** be μ-non-chaotic where μ is a computable measure. We have to prove that **a** is μ-predictable.

The winning strategy will be continuous, i.e. on kth move it will predict kth term of **a**. And it will satisfy the following property: let $V(s)$ denote Man's capital after the game according to the strategy against finite sequence s, then $V(s) = \nu(s)/\mu(s)$.

Assume that there is such a (computable) strategy. As **a** is ν-chaotic, the ratio $\gamma_A(\mathbf{a}|n)/\nu(\mathbf{a}|n)$ is bounded, whereas the ratio $\gamma_A(\mathbf{a}|n)/\mu(\mathbf{a}|n)$ is unbounded, therefore the ratio $\nu(\mathbf{a}|n)/\mu(\mathbf{a}|n)$ is unbounded, i.e. Man wins in μ-game against **a**.

It remains to define a computable strategy satisfying the above condition with initial capital $V_0 = \nu(\Lambda)/\mu(\Lambda)$. But this problem is already solved: as ν is a measure, the following equality holds:
$V(s)\mu(s) = V(s\hat{}0)\mu(s\hat{}0) + V(s\hat{}1)\mu(s\hat{}1)$. Then we apply the Proposition 6.2.5.

We would note that the constructed strategy depends only on μ and ν but not on **a**. The case when $\mu(\mathbf{a}|n) = 0$ is easy. □

Let us not forget that the problem 8.11 of existence of an unpredictable non-chaotic sequence is open.

The following theorem states that we cannot strengthen Theorem 9.7 by asserting that for natural sequences chaoticness is equivalent to stochasticness.

Theorem 9.8 (van Lambalgen–Shen–Muchnik). *There exists a stochastic natural sequence which is predictable (and hence non-chaotic) with respect to the uniform measure.* (Cf. [21, no. 6.2.4].)

Proof. We refer here to Corollary 3.14, Proposition 5.4, Theorems 8.7, 8.9, 5.1, 7.5. Those assertions imply the following. To get the required sequence it suffices to take any sequence which is chaotic relative to a quasiuniform measure with the parameter $\langle p_1, p_2, p_3 \ldots \rangle$ such that d_n tends to zero very slowly (where $d_n = p_n - 0.5$ as in Theorem 5.1). The details are left to the reader. □

Open question 9.9. Is there a computable strategy which wins in the game against all predictable natural sequences?

We conclude this section with the theorem related with the notion of lawlessness introduced in Section 2.

Theorem 9.10 (An.A. Muchnik). *All lawless sequences are predictable (even monotonically) with respect to every computable measure and, therefore, are not natural.*

Proof. We will consider first the case of the uniform measure η. The strategy which wins in η-game against any lawless sequence **a** is rather simple: on kth move it bets the half of current capital on value 0 of kth term of the seouence.

It is clear that after k moves the current capital is equal to $V_0(0.5)^{1(k)}(1.5)^{0(k)}$, where $0(k)$ and $1(k)$ denote the amount of zeros and ones, respectively, among k first terms of the sequence **a**. Therefore it is sufficient to prove that there are arbitrary large k such that all terms a_n, $k \leqslant n \leqslant 3k$, are equal to zero.

This follows from the lawlessness of **a**. Assume that this is not true; then for any sufficiently large k the segment $\mathbf{a}|k$ is continued in **a** by a sequence which has no prefix $000\ldots0$ ($2k$ times).

Now, let us consider general case. Let μ be a computable measure. The computable strategy is as follows. On the kth move it bets the half of current capital on the value i of kth term of the sequence **a**, where i is defined as follows. Let $s = \mathbf{a}|(k-1)$, then

$i=0$ if $\mu(s\hat{\ }0)\leqslant\mu(s\hat{\ }1)$ and $i=1$ else. Again there are arbitrary large k such that Man has at least $2k$ wins in succession after kth move (here we use computability of μ) and every win increases his capital at least 1.5 times. □

10. Finite sequences: unpredictability and chaoticness

From the point of view that is accepted in this paper only those mathematical assertions have a real meaning that deal only with finite objects. We consider infinite "sets" (if they are not euphemisms, for example, when speaking about infinite computable sequences rather than about algorithms computing them) as not more than a tool for really important analysis of finite objects (we mean a wide sense including the heuristics).

According to this concept, we consider the results of previous section as the results helping us to find a right way of analysis of the concept of randomness of *finite* objects in algorithmic probability theory.

Of course, in the finite case the notion of randomness has no absolute meaning. We have to find a way to measure the amount of randomness. A helpful tool for this measuring is the notion of complexity, as well as in the infinite case; but now we need another version of this notion – the *conditional entropy*.

Proceeding to the case of finite objects one should not only replace the unconditional entropy by the conditional one; one should also modify the very structure of the set Ξ. Up to here Ξ was treated as the set of finite initial segments of infinite sequences, and we had to know whether one of such segments is an extension of another segment. The structure of Ξ was a *tree* structure. This standpoint manifested itself in the additive clause of the definition of a quasimeasure in no. 3.3.4 and hence, implicitly, in Definition 8.1. Now, in the finite case, we deal with the elements of Ξ as with separate, discrete objects, paying no attention to the extension relation. So in this case the structure Ξ is a *bunch* structure (cf. [19, no. 1.3]). By this cause there is no additive requirement, which connects the values of $x, x\hat{\ }0, x\hat{\ }1$, in Definition 10.1 below.

Definition 10.1. A *conditional semimeasure* on the set $\mathbb{X}$ with the *space of conditions* $\mathbb{W}$ is any function $\nu:\mathbb{W}\times\mathbb{X}\to[0,+\infty)$ such that $\sum_{x\in\mathbb{X}}\nu_w(x)\leqslant 1$ for all $w\in\mathbb{W}$. (We write $\nu_w(x)$ instead of $\nu(w,x)$.) The elements of $\mathbb{W}$ are called *conditions*.

let $\mathbb{W}$ and $\mathbb{X}$ be constructive spaces. (A *constructive space* is a set with an effective bijection onto $\mathbb{N}$.) We give the following

Definition 10.2. A conditional semimeasure is *recursively enumerable from below* (r.e.b. in brief) if the "undergraph"

$$\{\langle w,x,q\rangle\in\mathbb{W}\times\mathbb{X}\times\mathbb{Q}\colon q<\nu_w(x)\}$$

is recursively enumerable.

Let $\mathbb{X}$ and $\mathbb{W}$ be fixed. Then, as in the case of semimeasures and with the similar quasiordering (see Section 8), the family of all conditional semimeasures recursively enumerable from below has the greatest conditional semimeasure. For every pair of constructive spaces $\mathbb{W}, \mathbb{X}$ we fix one of the greatest conditional recursively enumerable form below semimeasures on $\mathbb{X}$ with conditions from $\mathbb{W}$. This semimeasure is denoted by ρ. Thus, the value $\rho_w(x)$ is defined. (We suppose that the sets $\mathbb{W}$ and $\mathbb{X}$ are fixed together with bijections from those sets onto $\mathbb{N}$, this fixation will be done in every context. We have no space here to discuss the dependence of the definition on fixed bijections.) In particular, the notation $\rho_n(s)$ will mean that we are considering the conditional semimeasure on the set Ξ of all finite binary sequences with conditions from the set of natural numbers.

Definition 10.3. The semimeasure ρ is called the *greatest conditional semimeasure*. The complexity $K : \mathbb{W} \times \mathbb{X} \to \mathbb{N}$ defined by equality $K_w(x) = [-\mathrm{lb}(\rho_w(x)) + 1]$ is called *the conditional entropy*. In particular, we denote by K_n the conditional entropy when the condition is $n \in \mathbb{N}$.

For each kind of entropy (see [19 or 22]) there exists the conditional variant.

Thus, if ρ' is another r.e.b. conditional semimeasure of $\mathbb{X}$ with conditions from $\mathbb{W}$ then there is a positive constant c such that $\rho'_w(x) \leqslant c\rho_w(x)$ for all x, w.

The notation $\rho_{w_1,w_2}(x)$ (with a pair or a longer tuple of conditions) will mean that the set of conditions $\mathbb{W}$ is the Cartesian product $\mathbb{W}_1 \times \mathbb{W}_2$ and $w_i \in \mathbb{W}_i$. The meaning of notation K_{w_1,w_2} conforms to that.

We begin the analysis of finite case with the study of finite games and unpredictability. We suppose for convenience that initial Man's capital is equal to one. Suppose also that a computable quasimeasure μ on Ξ is fixed in sense of Definitions 3.3.4 and 3.3.6.

Now we are going to present an analog of the notion of predictability for finite sequences. However, at first we need to clarify some points concerned with game against a finite sequence. Let us fix some $n \in \mathbb{N}$ and consider the case when Man plays in the game "For cash" (described in Section 6.1) against a finite sequence $s \in \Xi$ of length n. Let us change the rules of that game as follows: Man can choose the terms for prediction in any order. The rule of changing capital is like the rule in infinite μ-game (see Section 6.2).

Definition 10.4. We will denote by $\mathbb{M}_n$ the set of all rational-valued probabilistic measures on the set I^n of binary sequences of length n. Thus, $\mathbb{M}_n$ consists of all functions defined on I^n with values in the set of positive rational numbers such that the sum of all values is equal to 1.

It is reasonable to take any measure $\mu \in \mathbb{M}_n$ as the measure presented in the rule of changing the capital. Let us call the resulting game the μ-game. Any Man's strategy in μ-game is a finite object and instead of the requirement of computability of a strategy we will measure its complexity (better to say, its entropy). In the finite case without loss of generality we may consider only totally defined and correct strategies that make exactly n moves in the game against any sequence of length n. The set of all such strategies is denoted by $\Sigma(n)$. Let us call Man's capital after the nth move in the game according to a strategy $\sigma \in \Sigma(n)$ against a sequence s of length n the *final capital* and let us denote it by $V^\sigma(s)$ (remind that the initial capital is equal to 1).

Definition 10.5. Let $\alpha, \beta, n \in \mathbb{N}$ and let $\mu \in \mathbb{M}_n$. A sequence $s \in I^n$ is called μ-α-β-predictable if there is a strategy σ such that $K_n(\sigma) < \alpha$ and $V^\sigma(s) \geqslant 2^\beta$ in the μ-game against s.

Let us explain why it is natural to compare the complexity of a strategy with the logarithm of amount of winnings. Both parameters α and β express in a sense the complexity or quantity of information. Assume, for example, that Man knows the values of β first terms of the sequence s ($\beta \leqslant n = \mathrm{lh}(s)$), i.e. Man possesses β bits of information about s, then he can make the final capital to be equal to 2^β in η-game against s. One can easily verify that the conditional entropy K_n of the meant strategy is about β (to within additive terms $2\,\mathrm{lb}(\beta) + C$). Thus, if Man possesses β bits of information about s then he can hope get the capital 2^β. Suppose that less information, for example, α bits, where $\alpha < \beta$, is enough to get the capital 2^β. This assumption restricts the class of sequences, and the exact formulation of this restriction is μ-α-β-predictability. In other words, if we have a strategy σ of complexity $K_n(\sigma) = \alpha$ and wish to win the capital 2^β (i.e. greater than naturally expected 2^α) we need $\beta - \alpha$ additional bits of information about s. The existence of such an information means μ-α-β-predictability.

As we see, the difference $\beta - \alpha$ is essential. Let us denote $\delta = \beta - \alpha$.

Definition 10.6. A sequence s of length n is called μ-δ-chaotic if $\rho_n(s) < 2^\delta \mu(s)$. A sequence which is not μ-δ-chaotic is called μ-δ-nonchaotic.

As in infinite case, the following theorem holds

Theorem 10.7. *Every μ-α-β-predictable sequence s of length n is μ-$(\beta - \alpha + C)$-nonchaotic, where C is a constant not depending on n, α, β and s.*

Proof. Let us define $v_n(t) = 0$ if $\mathrm{lh}(t) \neq n$ and $v_n(t) = \sum_{\sigma \in \Sigma(n)} \rho_n(\sigma) V^\sigma(t) \mu(t)$ if $\mathrm{lh}(t) = n$.

We claim that v is a recursively enumerable from below conditional semimeasure on Ξ. The recursively enumerability from below follows from the recursively enumerability from below ρ_n. Thus, it remains to verify that $\sum_s v_n(s) \leqslant 1$ for all n. We have

$$\sum_s v_n(s) = \sum_{\sigma \in \Sigma(n)} \rho_n(\sigma) \sum_{s \in I^n} V^\sigma(s) \mu(s).$$

But the inner sum is equal to 1 for any strategy σ (the analog of the Proposition 6.2.2). Hence $\sum_s \nu_n(s) = \sum_{\sigma\in\Sigma(n)} \rho_n(\sigma) \leqslant 1$, as ρ is a semimeasure.

Thus ν is really r.e.b. conditional semimeasure. Therefore, there is a constant c such that $\rho_n(s) \geqslant c\nu_n(s)$ for all n and s.

Let s be a μ-α-β-predictable sequence of length n and let $\sigma \in \Sigma(n)$ be a strategy such that $K_n(\sigma) < \alpha$ and $V^\sigma(s) \geqslant 2^\beta$. We obtain $\rho_n(s) \geqslant c\rho_n(\sigma)V^\sigma(s)\mu(s) \geqslant c2^{-\alpha}2^\beta\mu(s) = 2^{\beta-\alpha+C}\mu(s)$, where $C = \mathrm{lb}(c)$. □

Theorem 10.8. *The μ-measure of the set of all μ-δ-nonchaotic sequences of given length n (i.e. the sum of μ-measure of all elements from this set) is at most $2^{-\delta}$.*

Proof. If a sequence s is μ-δ-nonchaotic then $\rho_n(s) \geqslant \mu(s)2^\delta$. When s ranges the set of all μ-δ-nonchaotic sequences, the sum of all left hand sides of this inequality does not exceed 1 and the sum of right hand sides is equal to the μ-measure of this set multiplied by 2^δ. □

Now we give the last main definition of our paper.

Definition 10.9. A sequence s of length n is called *θ-δ-natural* if there is a measure $\mu \in \mathbb{M}_n$ such that s is μ-δ-chaotic and $K_{n\theta}(\mu) < \theta$. (Recall that the subscript $n\theta$ means the code of pair $\langle n, \theta\rangle$.) A not θ-δ-natural sequence is called *θ-δ-unnatural.*

Kolmogorov put the following question: are there (if θ and δ satisfy some reasonable requirements) θ-δ-unnatural sequences? Shen' [17] gave the positive answer, then V'yugin [25, Theorems 2 and 3] estimated the measure of the set of θ-δ-unnatural sequences. Muchnik improved that estimate and it got its final form. Let us denote, for the set $U \subseteq I^n$, by $\rho_w[U]$ the sum $\sum_{s\in U} \rho_w(s)$; the expression $\mu[U]$ for $\mu \in \mathbb{M}_n$ is understood in a similar way. We use this notation in order to make differences with the expression $\rho_w(U)$, which means a measure (i.e. the greatest conditional semimeasure) of U as a finite object. We denote by $U_{n\theta\delta}$ the set of all θ-δ-unnatural sequences of length n.

Theorem 10.10 (An.A. Muchnik). *For all sufficiently large positive integers n, θ, δ, if $\delta < n - 3\theta$ then $c_1 2^{-\theta} < \rho_{n\theta}[U_{n\theta\delta}] < c_2 2^{-\theta}$, where c_1, c_2 are positive constants not depending on n, θ, δ.*

Proof. The lower bound. In fact we will construct a *single* sequence $s \in U_{n\theta\delta}$ that will insure the inequality $\rho_{n\theta}[U_{n\theta\delta}] > 2^{-\theta}c_1$ and simultaneously we will define c_1. We will use a technical lemma which in spite of simplicity may have many other applications:

Lemma 10.10.1. If f is computable function then $K_w(f(w,x)) \leqslant K_w(x) + C$ *for all x from the domain of f_w, where C is a constant not depending on x, w. (We write $f_w(x)$ instead of $f(w,x)$.)*

Proof. Let us set $\gamma_w(y)=\sum_x \rho_w(x)$, where the sum is over the set of all x such that $f(w,x)=y$. It is easy to check that the function γ is a r.e.b. conditional semimeasure. Then $\rho_w(f(w,x))\geqslant c\gamma_w(f(w,x))\geqslant c\rho_w(x)$, which implies the desired inequality for $C=1-\mathrm{lb}(c)$. $\square$

Let us denote by $M_{n\theta}$ the set of all measures $\mu\in\mathbb{M}_n$ satisfying the inequality $\rho_{n\theta}(\mu)>2^{-\theta}$. The set $M_{n\theta}$ contains at most 2^{θ} measures. For $\mu\in M_{n\theta}$ we define $T_\mu=\{s\in I^n:\mu(s)>2^{\theta-n}\}$. Every T_μ has less than $2^{n-\theta}$ elements, therefore, there is a sequence of length n which belongs to no set T_μ, $\mu\in M_{n\theta}$. Let us denote by s_n the lexicographically least such sequence.

Lemma 10.10.2. $\rho_{n\theta}(s_n)>c_1 2^{-\theta}$, *where* c_1 *is a constant not depending on* n *and* θ.

Proof. Let $\mathbb{M}=\bigcup_{n\in\mathbb{N}}\mathbb{M}_n$. As ρ is recursively enumerable from below, there is a computable function $h:\mathbb{N}\to\mathbb{N}\times\mathbb{M}\times\mathbb{Q}$ whose range coincides with the "undergraph" of the semimeasure ρ. For $\mu\in M_{n\theta}$ we denote by i_μ the least i such that $h(i)=\langle\langle n,\theta\rangle,\mu,r\rangle$, for some $r>2^{-\theta}$. One can easily verify that $i_\mu\neq i_\nu$ if $\mu\neq\nu$. Let us take the measure $\mu\in M_{n\theta}$ with maximal value of i_μ.

Given μ,n,θ we can compute $M_{n\theta}$ as a finite object (using h we at first find the number i_μ and then using the values $h(i)$, $i\leqslant i_\mu$ we find all other measures from $M_{n\theta}$). On the other hand, given n,θ and $M_{n\theta}$ we can find s_n. Applying Lemma 10.10.1 we get $\rho_{n\theta}(s_n)\geqslant\rho_{n\theta}(\mu)c$ for a constant c not depending on n,θ. As $\rho_{n\theta}(\mu)>2^{-\theta}$, we get $\rho_{n\theta}(s_n)\geqslant c_1 2^{-\theta}$. $\square$

Thus, it remains to prove that $s_n\in U_{n\theta\delta}$. Let a measure $\mu\in\mathbb{M}_n$ satisfy the inequality $K_{n\theta}(\mu)<\theta$, that is $\mu\in M_{n\theta}$. We have to prove that s_n is not μ-δ-chaotic, i.e. that $\rho_n(s_n)\geqslant 2^{\delta}\mu(s_n)$. Because of the choice of s_n we have $s_n\notin T_\mu$, that is $\mu(s_n)\leqslant 2^{\theta-n}$.

On the other hand, we have proved that $\rho_{n\theta}(s_n)>2^{-\theta}c$ for a constant c not depending on n and θ. Let us use the auxiliary inequality: $\rho_n(s)\geqslant c'\theta^{-2}\rho_{n\theta}(s)$, where c' is a constant not depending on n,θ and s. (Proof: the function $\nu_n(s)=0.5\sum_\theta\theta^{-2}\rho_{n\theta}(s)$ is recursively enumerable from below and is majorized by a measure since $\sum i^{-2}<2$.) Hence $\rho_n(s_n)>c\theta^{-2}2^{-\theta}$, where c is a (new) constant not depending on n,θ.

As $\mu(s_n)\leqslant 2^{\theta-n}$, we get $\rho_n(s_n)>c\mu(s_n)2^{n-2\theta}\theta^{-2}$. If θ is sufficiently large then the last inequality implies $\rho_n(s_n)>\mu(s_n)2^{n-3\theta}$. If $\delta\leqslant n-3\theta$ then we obtain $\rho_n(s_n)>\mu(s_n)2^{\delta}$.

Now, the upper bound of Theorem 10.10. It is sufficient to construct a measure $\mu\in M_{n\theta}$ such that $\rho_{n\theta}[H_{\mu\delta}]<2^{-\theta}c_2$, where $H_{\mu\delta}$ is the set of all μ-δ-nonchaotic sequences (of length n) and c_2 is a constant.

This measure μ is constructed as follows. Let a number $k\leqslant\theta$ be fixed. Let us denote by $r=r_{n\theta k}$ the greatest rational number with denominator $2^{\theta-k}$ which is less than the sum $\sum_{s\in I^n}\rho_{n\theta}(s)$. We claim that given r,n,θ and k we can compute a semimeasure $\nu=\nu_{n\theta k}$ on I^n such that $\nu(s)\leqslant\rho_{n\theta}(s)$ for all $s\in I^n$ and $\sum_{s\in I^n}\nu(s)\geqslant r$. Indeed, let us execute the procedure of generating of all pairs $\langle q,s\rangle$ such that $s\in I^n$, $q\in\mathbb{Q}$, $q<\rho_{n\theta}(s)$. Let $i=i_r$ be the first step on which it is generated many enough pairs to prove that

$\sum_{s\in I^n} \rho_{n\theta}(s) \geqslant r$. Take as $\nu = \nu_{n\theta k}$ the obtained "lower" approximation of $\rho_{n\theta}$. Finally, take as $\mu = \mu_{n\theta k}$ the measure from $\mathbb{M}_n$ obtained from ν by increasing the value $\nu(s_0)$, where s_0 is the sequence consisting of n zeros, to obtain $\sum_{s\in I^n} \mu(s) = 1$. So, μ (like ν) can be computed given r, n, θ and k.

Applying the analog of Lemma 10.10.1 we get $\rho_{n\theta k}(\mu) \geqslant c\rho_{n\theta k}(r)$, where c is a constant not depending on n, θ, k and r. As r is in fact a binary sequence of length θ–k we have $\rho_{n\theta k}(r) \geqslant c_3 2^{k-\theta}$ for some constant c_3 (to prove this we have to consider the measure $\varphi_{n\theta k}(s) = 2^{k-\theta}$ if $s \in I^{\theta-k}$ and $\varphi_{n\theta k}(s) = 0$ else). As in the proof of the lower bound, we can prove that $\rho_{n\theta}(\mu_{n\theta k}) \geqslant c' k^{-2} \rho_{n\theta k}(\mu_{n\theta k})$. Thus, from three last inequalities, we conclude that $\rho_{n\theta}(\mu_{n\theta k}) \geqslant c_4 k^{-2} 2^k 2^{-\theta}$ where c_4 is a constant not depending on n, θ, and k. For sufficiently large $k \leqslant \theta$ we get $\rho_{n\theta}(\mu_{n\theta k}) > 2^{-\theta}$.

Thus, it remains to verify that $\rho_{n\theta}[H_{\mu\delta}] < c_2 2^{-\theta}$.

Note that if $s \in H_{\mu\delta}$ then $\rho_n(s) \geqslant \mu(s)2^\delta$. Clearly, $\rho_{n\theta}(s) \geqslant c''\rho_n(s)$. Therefore, for every $s \in H_{\mu\delta}$ the inequality $\rho_{n\theta}(s) \geqslant \mu(s)2^\delta c''$ holds. If δ is sufficiently large then the last inequality implies $\rho_{n\theta}(s) \geqslant 2\mu(s)$. Therefore, $\rho_{n\theta}[H_{\mu\delta}] \geqslant 2\mu[H_{\mu\delta}]$, or $\rho_{n\theta}[H_{\mu\delta}] - \mu[H_{\mu\delta}] \geqslant \rho_{n\theta}[H_{\mu\delta}]/2$.

By construction of ν and μ the left hand side of the last inequality is at most $2^{k-\theta}$. Therefore, $\rho_{n\theta}[H_{\mu\delta}] \leqslant 2^{k+1} 2^{-\theta}$. As the choice of k does not depend on n, θ, δ we can define $c_2 = 2^{k+2}$. □

We conclude this section with the discussion of an another possible definition of chaoticness and naturalness. Namely, let us say that a binary sequence s of length n is *μ-δ-chaotic in the second sense* if $\rho_{n\mu}(s) < 2^\delta \mu(s)$ (the difference with Definition 10.6 is that μ is added as an additional condition). The definition of *θ-δ-naturalness in the second sense* is obtained from the Definition 10.9 by replacing μ-δ-chaoticness with μ-δ-chaoticness in the second sense.

What version is more reasonable? We think that if the measure μ is given and we estimate the probability of appearing of sequences, then it is reasonable to relativize the a priori probability with the measure (the second version). Conversely, if we already have a sequence and search for a simple measure relative to which our sequence is random then it is reasonable to relativize the a priori semimeasure with the length but not with the measure (the first version).

Theorem 10.10 for the new definition becomes the following Theorem 10.11:

Theorem 10.11 (An.A. Muchnik). *For all sufficiently large positive integers n, θ, δ, if $\theta < \delta < n - 3\theta$ then $c_1 2^{-\theta} \leqslant \rho_{n\theta}[U'_{n\theta\delta}] \leqslant c_2 2^{-\theta}$, where $U'_{n\theta\delta}$ is the set of all θ-δ-unnatural in the second sense sequences and c_1, c_2 are constants not depending on n, θ, δ.*

Proof. The lower bound can be obtained from the proof of the lower bound of Theorem 10.10 as follows. We have to prove that $\rho_{n\mu}(s_n) \geqslant 2^\delta \mu(s_n)$ (we use the notation introduced before Lemma 10.10.2). Let us make use of the proven inequality:

$$\rho_n(s_n) > c\mu(s_n) 2^{n-2\theta} \theta^{-2}.$$

On the other hand, $\rho_{n\mu}(s_n) \geqslant c'\rho_n(s_n)$ where c' does not depend on n and μ. Hence, if θ is sufficiently large and $\delta < n - 3\theta$, then the inequality $\rho_{n\mu}(s_n) \geqslant 2^\delta \mu(s_n)$ holds.

Let us turn to the upper bound. We denote by $H'_{\mu\delta}$ the set of all sequences (of length n) μ-δ-nonchaotic in the second sense. We want to prove that $\rho_{n\theta}[H'_{\mu\delta}] < 2^{-\theta}c$ for the measure μ constructed in the proof of the upper bound of Theorem 10.10 (remind that μ depends on the parameter k).

We have $\rho_{n\mu}(s) \geqslant 2^\delta \mu(s)$ for all $s \in H'_{\mu\delta}$ and we want to replace in this inequality the subscript μ by the subscript θ.

Thus, we have

$$\frac{\rho_{n\mu}(s)}{\mu(s)} = \frac{\rho_{n\mu}(s)}{\rho_{n\mu\theta}(s)} \frac{\rho_{n\mu\theta}(s)}{\rho_{n\theta}(s)} \frac{\rho_{n\theta}(s)}{\mu(s)} \geqslant 2^\delta.$$

Let us denote by X, Y and Z the multipliers in the left hand side of last inequality. Then $X \leqslant c'$ for a constant c' not depending on n, θ, μ, and s. Secondly, we consider the set H_1 of all s for which $Z \geqslant 2$. By construction of the measure μ $\sum_{s\in I^n}(\rho_{n\theta}(s) - \mu(s)) \leqslant 2^{k-\theta}$. Hence $\rho_{n\theta}[H_1] \leqslant 2^{k+1}2^{-\theta}$. Thirdly, we consider the set H_2 of all s for which $Y \geqslant (c')^{-1}2^{\delta-1}$. We see that $\rho_{n\theta}[H_2] \leqslant c'2^{1-\delta}\rho_{n\mu\theta}[H_2] \leqslant 2c'2^{-\delta}$ and for $\delta > \theta$ we have $\rho_{n\theta}[H_2] \leqslant 2c'2^{-\theta}$.

However, it is clear that $H'_{\mu\delta} \subseteq H_1 \cup H_2$. Thus, we can take $c_2 = 2c' + 2^{k+1}$. (In the previous proof the use of $\rho_{n\mu\theta}$ is not necessary, but in other cases it can be useful.) $\square$

11. Open questions

For the convenience of the reader we remind here the mathematical problems having been arisen in our paper.

1. Are the notions of chaoticness and unpredictability equivalent? [5]

2. Is there a computable strategy which wins in the game against all predictable natural sequences?

12. A philosophical supplement

The discussion of philosophical problems related with the exposed results began in the Introduction. Let us return to that discussion.

As it was mentioned in the Introduction, the appearance of the tuple consisting of 20 zeros is surprising and the appearance of the tuple say 01111011001101110001 is not surprising. Now, when the mathematical basis of the theory has been formulated in Sections 1–10, we can try to explain, in more formal terms, the causes of our surprise or its absence.

[5] As it is clear, the definitions of these notions are of different topological nature: chaoticness, by definition, is Σ_2-formula ($\exists\forall$) whereas unpredictability is Π_3-formula ($\forall$ strategy $\exists$ bound of capital $\forall$ number of a move (current capital does not exceed the bound)).

As it was mentioned in the commentary to the Definition 8.4, the a priori entropy, as a function of a tuple, is defined to within an additive constant term; to say more precisely: to within an additive bounded function. Thus, the final choice of a function to be the entropy is arbitrary within the said limits; therefore, it is meaningless to speak on the value of the entropy of an individual object. However, one may hope that there are more natural as well as less natural entropies, and we will admit this assumption. Though no natural mathematical definition of what is a good, or natural, entropy function is known to the authors, the authors nevertheless believe that it is possible, in theory, to distinguish good (natural) entropy functions. Presuming that this distinguishing is possible, we can consider that the tuple consisting of only zeros has small entropy and therefore has large complexity deficiency (provided the tuple is sufficiently large). Let us remind that the complexity deficiency is the length of a tuple minus its a priori entropy.

We think that the cause of surprise in the above example is the amount of complexity deficiency. Namely, we are surprised if that amount is large.

Remark 1. The above statement is true only for the uniform distribution; however, it can be generalized to an arbitrary case. In the general case, we have to replace the complexity deficiency with the randomness deficiency. The randomness deficiency of a tuple is the absolute value of the logarithm of its probability minus its a priori entropy. In the uniform case, the absolute value of the logarithm of the probability becomes the length and the randomness deficiency becomes the complexity deficiency.

Our discussion of the reasons to qualify an event as surprising have direct connections with basic problems of the mathematical statistics. One of those problems is as follows. Let a finite sequence of zeros and ones, or a tuple, appear in an experiment; the following problem arises: to find a probability distribution with respect to which the sequence is not surprising. According to our agreement, we identify the surprise with a large amount of the randomness deficiency. Then the problem becomes to have an exact mathematical meaning (let us remind that we fixed some particular a priori entropy).

In this refined form, the problem was posed by A.N. Kolmogorov in the 1960s and 1970s before participants of his seminar at Moscow State University. Kolmogorov accompanied his refinement of the statistical problem by the following two natural questions – the question whether the required distribution is unique and the question whether it exists. It is rather obvious that the required distribution is, in general case, not unique. It turned out, that the required distribution may not exist, this was discovered by Shen' ([17]; cf. [8, Section 2.5], [9, Section 2.5]). The question whether this state of affairs is real belongs to philosophy and to natural science but not to mathematics.

The question is concerned with the structure of the real world and consists in the following. Is there a sequence in the real world which has no adequate probability distribution, i.e., no distribution with respect to which the sequence is not surprising? We can state as a hypothesis that such a sequence does not exist – in the reality, not in the world of mathematical abstractions. In other words, in the real physical world, there are no "inherently surprising" sequences, i.e., sequences that remain surprising under any changes of probabilistic assumptions. This hypothesis is ascribed sometimes to Kolmogorov; however, the authorship by Kolmogorov is doubtful. It would be more right to say that hypothesis was discussed among participants of the Kolmogorov seminar, therefore we will call it "Kolmogorov Seminar Hypothesis".

Remark 2. In the case of infinite sequences the "Kolmogorov Seminar Hypothesis" looks as follows: any infinite sequence present in the nature is chaotic with respect to some computable probability distribution, i.e., is natural in the sense of Definition 9.6. Thus the hypothesis warrants the use of the term "natural" in that definition. Of course, we abstract ourselves from the fact that infinite sequences do not exist in the nature at all.

The following consideration may be opposed to "Kolmogorov Seminar Hypothesis". Let us write down *all* the finite sequences in succession, then inherently surprising sequences will appear some time or other: indeed, their existence was proved by Shen'. However, the getting of any individual finite sequence requires, in this process, very large amount of time – exponential of the length of the sequence. We can consider that time greater than the time of existence of the Universe in which the "Kolmogorov Seminar Hypothesis" is true.

On the other hand, the following finite analog of Gács' Theorem (i.e., Theorem 8.8) is true (and can be proved simpler than the infinite analog):

any finite sequence can be obtained as the result of certain simple (i.e. having a simple description) transformation of an appropriate chaotic finite sequence having the length approximately equal to the entropy of the given sequence.

It may seem that this finite analog makes inconsistent

(1) the Kolmogorov Seminar Hypothesis and

(2) the identifying the surprise with the large amount of complexity deficiency. Indeed, we should consider the chaotic finite sequence which, as the result of a simple transformation, the surprising finite sequence can be obtained from, as surprising, too; but that sequence cannot be surprising since it has small complexity deficiency (as it is chaotic).

However, the inconsistency is seeming. The transformation (though having a simple description) can require very large amount of time. And we have already pointed out the crucial role of time of process durations.

Acknowledgements

The authors are thankful to Nikolai K. Vereshchagin for translating the Russian original, and to Konstantin Yu. Gorbunov for his help in preparing the document.

References

[1] P.J. Cohen, Set Theory and the Continuum Hypothesis, W.A. Benjamin, New York, 1966.

[2] P. Gács, Exact expressions for some randomness tests, Zeitschr. f. math. Logik und Grundlagen d. Math. 26 (1980) 385–394.
[3] P. Gács, Every sequence is reducible to a random one, Inform. and Control 70 (2/3) (1986) 186–192.
[4] S. Kakutani, On the equivalence of infinite product measures, Ann. Math. 49 (1948) 214–224.
[5] A.N. Kolmogorov, On tables of random numbers, Sankhyā. The Indian J. Statist. Ser. A. 25 (4) (1963) 369–376.
[6] A.N. Kolmogorov, Logical basis for information theory and probability theory, IEEE Trans. Inform. Theory IT-14 (1968) 662–664.
[7] A.N. Kolmogorov, On the logical foundations of information theory and probability theory, Problems of Inform. Transmission. 5 (3) (1969) 1–4.
[8] A.N. Kolmogorov, V.A. Uspensky, Algorithms and randomness, Proc. 1st World Congr. of the Bernoulli Soc. Tashkent, 1986. Vol 1: Probability Theory and Applications, VNU Science Press, Utrecht, 1987, pp. 3–53.
[9] A.N. Kolmogorov, V.A. Uspensky, Algorithms and randomness, Theory of Probability and its Applications 32 (1987) 389–412.
[10] L.A. Levin, On the notion of a random sequence, Soviet Math. Doklady 14 (1973) 1413–1416.
[11] P. Martin–Löf, The definition of random sequences, Inform. and Control 9 (6) (1966) 602–619.
[12] P. Martin–Löf, On the notion of randomness, in: A. Kino, J. Myhill, R.E. Vesley (Eds.), Intuitionism and Proof Theory, New York, 1968, pp. 73–78.
[13] P. Martin–Löf, Notes on Constructive Mathematics, Almqvist & Wiksell, Stockholm, 1970.
[14] R. von Mises, Grundlagen der Wahrscheinlichkeitsrechnung, Mathematische Zeitschrift 5 (1919) 52–89.
[15] R. von Mises, Wahrscheinlichkeit, Statistik und Wahrheit, Springer, Wien, 1928.
[16] R. von Mises, Probability, Statistics and Truth, 2d rev., English ed., Prep by H. Geiringer, Allen London; Macmillan, New York, 1957.
[17] A.Kh. Shen', The concept of (α, β)-stochasticity in the Kolmogorov sense, and its properties, Soviet Math. Dokl. 28 (1) (1983) 295–299.
[18] C.P. Schnorr, Process complexity and effective random tests, J. Comput. System Sci. 7 (1973) 376–388.
[19] V.A. Uspensky, Complexity and entropy: an introduction to the theory of Kolmogorov complexity, in: Watanabe (Ed.), Kolmogorov Complexity and Computational Complexity, Springer, Berlin, 1992, 85–102.
[20] V.A. Uspensky, A.L. Semenov, Theory of Algorithms: Main Discoveries and Applications, "Nauka", Moscow, 1987 (Russian). English translation titled "Algorithms: Main Ideas and Applications" is published by Kluwer Academic Publishers, Dordrecht, 1993.
[21] V.A. Uspensky, A.L. Semenov, A.Kh. Shen', Can an individual sequence of zeros and ones be random?, Russian Math. Surveys 45 (1) (1990) 121–189.
[22] V.A. Uspensky, A.Kh. Shen', Relations between varieties of Kolmogorov complexities, Math. Systems Theory 29 (1996) 271–292. (Also CWI Tech. Rep. CS-R9329.)
[23] V.V. V'yugin, Algorithmic entropy (complexity) of finite objects and its applications to defining randomness and amount of information, Semiotica i Informatika 16 VINITI, Moscow (1981) 14–43 (in Russian).
[24] V.V. V'yugin, Algorithmic entropy (complexity) of finite objects and its applications to defining randomness and amount of information, Selecta Mathematica formerly Sovietica 13 (4) (1994) 357–389.
[25] V.V. V'yugin, Nonstochastic objects, Problems of Information Transmission 21 (2) (1985) 77–83.
[26] A. Wald, Sur la notion de collectif dans le calcul des probabilités, C. R. Acad. Sci. Paris. 202 (1936) 180–183.
[27] A. Wald, Die Widerspruchsfreiheit des Kollektivbegriffes der Wahrscheinlichkeitsrechnung, Ergebnisse eines Math. Kolloquiuns Vienna 8 (1937) 38–72.
[28] A. Wald, Die widerspruchsfreiheit des kollektivbegriffes, Actualités Sci. Indust. 735 (1938) 79–99.
[29] O. Watanabe (Ed.), Kolmogorov Complexity and Computational Complexity, Springer, Berlin, 1992.
[30] A.K. Zvonkin, L.A. Levin, The complexity of finite objects and the development of the concepts of information and randomness by means of the theory of algorithms, Russian Math. Surveys 25 (6) (1970) 83–124.