ISSN 0032-9460, Problems of Information Transmission, 2007, Vol. 43, No. 1, pp. 48–56. © Pleiades Publishing, Inc., 2007.
Original Russian Text © K.Yu. Gorbunov, 2007, published in Problemy Peredachi Informatsii, 2007, Vol. 43, No. 1, pp. 56–66.

LARGE SYSTEMS

Bound on the Cardinality of a Covering of an Arbitrary Randomness Test by Frequency Tests[1]

K. Yu. Gorbunov

Kharkevich Institute for Information Transmission Problems, RAS, Moscow
gorbunov@iitp.ru

Received October 17, 2006

Abstract—We improve a well-known asymptotic bound on the number of monotonic selection rules for covering of an arbitrary randomness test by frequency tests. More precisely, we prove that, for any set S (arbitrary test) of binary sequences of sufficiently large length L, where $|S| \le 2^{L(1-\delta)}$, for sufficiently small δ there exists a polynomial (in $1/\delta$) set of monotonic selection rules (frequency tests) which guarantee that, for each sequence $\boldsymbol{t} \in S$, a subsequence can be selected such that the product of its length by the squared deviation of the fraction of zeros in it from 1/2 is of the order of at least $0.5 \ln 2\, L[\delta/\ln(1/\delta)](1 - 2\ln\ln(1/\delta)/\ln(1/\delta))$.

DOI: 10.1134/S0032946007010061

This paper answers a question stated in [1, Section 2.4]. Let us recall basic definitions and problem settings from [1] related to the considered problem. The problem itself will be formulated below.

An *arbitrary frequency test* for binary sequences of length L is any set S of such sequences. The *specific deficiency* $\delta(S)$ *of* S is $(L - \mathrm{lb}\,|S|)/L$, where lb is the binary logarithm and $|S|$ is the cardinality of S. Let $\boldsymbol{t}$ denote an arbitrary element of S. A *monotonic selection rule* is any selection rule (i.e., a function defined on all finite sequences and taking values 0 and 1) for selecting a subsequence of $\boldsymbol{t}$ which, at each step i ($i = 1, \ldots, L$), decides (using only the head of $\boldsymbol{t}$ of length $i-1$) whether the ith symbol t_i of $\boldsymbol{t}$ should be appended to the subsequence under construction. A *nonmonotonic* selection rule differs from a monotonic one in the following way: it may look through bits of $\boldsymbol{t}$ in an arbitrary order (the decision on appending a bit to the sequence under construction is made straight before reading this bit, based on the values of previously examined bits). The *specific deficiency of a rule conditional to* S is the number $\dfrac{2\,\mathrm{lb}\,e}{L}$ multiplied by the minimum (over all $\boldsymbol{t} \in S$) of the product of the length of a selected subsequence and the squared deviation of the fraction of zeros in it from 1/2 (here e is the base of the natural logarithm).

Let us briefly recall the reason for this definition (see details in [1]). A *normal rule* is a rule that always selects from a sequence of length L subsequences of the same length. Given any rule r, one can easily construct L normal rules $r_1, \ldots, r_L$, where each r_i selects a subsequence of "its own" length i and does not change the selection given by r in the cases where r also selects a subsequence of length i. To each normal rule r_i and a given deviation ε there corresponds a set T (*frequency test*) of subsequences of length L for which r_i selects a subsequence (of length i) with deviation of the fraction of zeros from 1/2 of at least ε. Using the bound for the large deviation probability (see [2, p. 93]; sometimes it is called the Chernoff bound), it is easy to show that the specific

[1] Supported in part by the Russian Foundation for Basic Research, project no. 06-01-00122, and the Grant of the President of the Russian Federation for State Support of Leading Scientific Schools, no. NSh-358.2003.1.

deficiency of T is at least $(2n\varepsilon^2 \operatorname{lb} e)/L$; if ε is small enough and n is large enough as compared to $1/\varepsilon$, then $\delta(T) \approx (2n\varepsilon^2 \operatorname{lb} e)/L$.

We say that a family R of rules δ'-*covers* S if there exists a partition of S into $|R|$ subsets and a one-to-one correspondence between the rules and subsets such that the specific deficiency of each rule conditional to the corresponding subset is larger than δ'. Note that, if the rules are normal, the meaning of this definition is clear: an arbitrary test S is covered by a family of $|R|$ frequency tests, and δ' indicates the "quality" of the frequency tests. Before formulating the main result, let us consider several examples and briefly describe known facts and open problems.

Example 1. Let L be even and assume that all positions are divided into pairs of neighboring positions. Consider the set S characterized by the following property: the number of pairs with two zeros is at most $0.5L(0.25 - \varepsilon)$, i.e., the fraction of such pairs differs by ε from the "proper" one quarter, where $\varepsilon < 0.1$. Clearly, the specific deficiency of S is larger than some positive number that depends on ε but is independent of L. Consider two monotonic rules: the first selects the first bits of all pairs to the subsequence, and the second looks at the first bit of each pair and, if it is 0, selects the second bit. It is easily seen that either the first rule attains the deviation of the fraction of zeros in the selected subsequence from 1/2 of at least ε, or so does the second rule; in the latter case, the length of the subsequence will be at least $L(0.5 - \varepsilon)$. Hence, the specific deficiency of at least one of the rules is not less than $(0.8 \operatorname{lb} e)\varepsilon^2$; i.e., the deficiency depends only on δ but not on L. It is clear that the same effect can be attained by using one nonmonotonic rule which first selects all first bits and then either stops (if the subsequence thus selected has large enough deviation of the fraction of zeros from 1/2) or selects the second bits in the pairs where the first bit is 0.

Example 2. Assume that all L positions are divided into blocks of even length d (the last block can be of length less than d; for fixed d and large L, this has no influence on the result). Consider the set S characterized by the property that the number of zeros in each block is equal to the number of ones. It is easily seen that the specific deficiency of S depends only on d but not on L (it is of the order of $0.5(\ln d)/d$). Consider two monotonic rules: the first rule selects the last bit of a block only if it is 0, and the second rule, only if it is 1 (it is clear that, when all bits except for the last are already read, the rule "understands" what the last bit is). Obviously, one of the rules attains the specific deficiency of approximately $(\operatorname{lb} e)/d$. It is easily seen that the same effect can be achieved by using one nonmonotonic rule which first finds out what the last bit of each block is and then selects the bits that form the majority.

A situation where one nonmonotonic rule is "sufficient," whereas several monotonic rules are necessary, is quite typical. In particular, the author is unaware of any case where one nonmonotonic rule is "insufficient"; more precisely, whether there exists some $\delta > 0$ such that, for any ε and any arbitrarily large L, there exists a set S of specific deficiency at least δ for which there is no nonmonotonic rule with specific deficiency conditional to S not less than ε. There is only known a corresponding statement in the case where ε is sufficiently large as compared with δ. More precisely, it is proved in [1, Theorem 5] that, if we require ε to be not less than $\dfrac{2\delta}{\ln \delta}$, then there exists S for which even some exponential (in L) number of nonmonotonic rules is insufficient.

For monotonic rules, the situation with lower bounds (i.e., with proofs that rules are "insufficient") is a little better. Thus, in the spirit of a well-known example by Ville (see [3, Section 6.2.2]), one can easily construct an example where one monotonic rule is "insufficient." As in Example 1, let all positions be divided into pairs. Consider a set S characterized by the property that there is no pair with two zeros. Then, given any monotonic rule, one constructs the following sequence: each time when the rule does not select the current bit to a subsequence, put it to be 1, and in the bits that are selected to the subsequence, let the values 0 and 1 alternate. It is clear that a sequence thus constructed belongs to S and that the rule does not attain any significant specific deficiency on it.

One can strengthen this lower bound and show that there are cases where of the order of $\ln(1/\delta)$ monotonic rules are "insufficient." As in Example 2, let all L positions be divided into blocks of the same length d, where d equals 2^{2r}, r being the number of rules that we want to "deceive." Consider a set S characterized by the property that, in each block, the number of zeros differs from the number of ones by at most 2^r. It follows from the central limit theorem that the specific deficiency, δ, of S equals c/d for some constant c; hence, $r = 0.5(\mathrm{lb}(1/\delta) + \mathrm{lb}\, c)$. Let there be r monotonic rules. Given them, we construct the following subsequence $\boldsymbol{t}$. Let i be the number of a current bit, and let M_i be the set of rules that decide to select this bit to a subsequence. If the set M_i did not occur earlier (i.e., if $M_i \neq M_j$ for each $j < i$), then we put 1 at the ith position; otherwise, consider the largest j for which $M_i = M_j$ and, at the ith position, put the bit other than what was put at the jth position (i.e., alternate bits corresponding to the same set of rules). Obviously, $\boldsymbol{t} \in S$. On the other hand, each rule deciding to select the next bit is "kept company" by one of the 2^{r-1} sets of rules, so that the number of zeros in the selected subsequence differs from the number of ones by at most 2^{r-1}. Clearly, its specific deficiency tends to zero as $L \to \infty$.

Known upper bounds are also related to monotonic rules (the author does not know whether they can be improved by allowing nonmonotonic rules). One of the first nontrivial upper bounds was obtained by An.A. Muchnik and is as follows (since a detailed proof of this theorem was never published (see [1, p. 164] and [4, Section 9.2.1]), we present it at the end of the paper).

Theorem 1. *For an arbitrary set S with specific deficiency δ there exist $c_1\dfrac{1}{\delta}$ monotonic rules that $c_2\delta^2$-cover S (here c_1 and c_2 are some constants).*

Now let us state our main result. It is proved in [1, Theorem 4] that, for any small enough $\delta > 0$, any $L \geq (1/\delta)^5$, and any set S of sequences of length L with specific deficiency δ, there exists a family, R, of monotonic rule that δ'-covers S, where $\delta' = \dfrac{\delta}{\ln(1/\delta)}(1-\beta)$, and for β one may take $\dfrac{2\ln\ln(1/\delta)}{\ln(1/\delta)}$. The cardinality of R is of the order of an exponent of $1/\delta$ (we mean not necessarily normal rules). The following theorem improves the bound on the cardinality of R. If $\delta' = \dfrac{\delta}{\ln(1/\delta)}(1-\beta)$, then, for fixed β, this cardinality becomes polynomial in $1/\delta$; for $\beta = \dfrac{2\ln\ln(1/\delta)}{\ln(1/\delta)}$, it is subexponential in $1/\delta$.

Theorem 2. *Let $\delta \in (0, \exp_e(-e^{50}))$, and let a natural $L \geq (1/\delta)^5$ be fixed. Consider sets of binary sequences of length L. For an arbitrary set S with specific deficiency of at least δ, there exists a family of at most*

$$\left(\frac{1}{\delta}\right)^{\frac{0.6\ln 2\ln(1/\delta)}{\beta\ln(1/\delta)-1.3\ln\ln(1/\delta)}+6.8}$$

monotonic rules that δ'-covers S, where $\delta' = \dfrac{\delta}{\ln(1/\delta)}(1-\beta)$. For β, one may take $\dfrac{2\ln\ln(1/\delta)}{\ln(1/\delta)}$.

Proof. In [1] (see the beginning of the proof of Theorem 4) there is considered a game where Mathematician and Nature make L moves in turn: at the ith move, Mathematician makes a bet of $x_i \in [0,1]$ on either 0 or 1, and Nature chooses an element $t_i \in \{0,1\}$ so that, after the Lth move, the constructed sequence $\boldsymbol{t}$ would belong to S. Initially, Mathematician's capital is zero; then, at each step, it increases by the bet if Mathematician guesses the next digit t_i and decreases by the same value otherwise (the capital can be negative as well). It is proved there that, for any set S, there exists a strategy of Mathematician that allows him to win at least $L\delta(S)\ln 2$. The bet at the ith step is $2\dfrac{|S_1(i)|}{|S_1(i)|+|S_2(i)|}-1$, where $S_1(i)$ and $S_2(i)$ are two sets of possible extensions of the known (before the ith move) heads of the sequence $\boldsymbol{t}$: one set with $t_i = 0$, and the other with $t_i = 1$; moreover, $|S_1(i)| \geq |S_2(i)|$ (Mathematician makes a bet on the bit that corresponds to $S_1(i)$).

Throughout what follows, when speaking about a game, we always assume that bets are computed according to this rule.

In the sequel, when speaking about probabilities, we mean the probabilities of events generated by the uniform distribution of the sought-for element $\boldsymbol{t}$ over the set S. We estimate the probabilities with the help of the following random process corresponding to the above-mentioned game. Namely, playing with Nature, Mathematician wins at the ith move with probability $\dfrac{|S_1(i)|}{|S_1(i)|+|S_2(i)|}$ and looses, respectively, with probability $\dfrac{|S_2(i)|}{|S_1(i)|+|S_2(i)|}$. Clearly, any sequence of S can equiprobably appear as the sequence $\boldsymbol{t}$ constructed at the end of the game.

Lemma 1. *Let M be a subset of the segment $[0,1]$, and let a and b be its lower and upper boundaries, respectively. Assume that n is the number of bets that belong to M, and let d be the difference between the number of wins and losses (for the considered n moves). Then, for any $\varepsilon \in (0, 0.1]$, if $L > 20/\varepsilon^4$ and $n \geq L\varepsilon$, the probability that $d < n(a-\varepsilon)$ is not greater than $\exp_e(-0.4L\varepsilon^3)$. The same is the estimate for the probability that $d > n(b+\varepsilon)$.*

Proof. Consider a binary tree T corresponding to the described random process (we represent it as "growing upwards"). Let us mark its vertices (except for leaves) by bets at them, which uniquely determine the probabilities of movements along edges. Vertices whose bets belong to M are refereed to as *active*. The *active height* of a vertex v is the number of active vertices along the way from the root to v (the vertex v itself is not counted). To estimate the probability of the first event in the statement of the lemma, for an arbitrary fixed n consider the following event $A(n,\varepsilon)$: along the way from the root to a leaf there are precisely n vertices and $d < n(a-\varepsilon)$. Let us modify the tree T as follows. Formally extend each leaf of active height less than n by a subtree (growing upward) so that new leaves have active height n, and make each vertex of active length n a leaf by deleting everything above it. Clearly, for the new tree T', the probability $A(n,\varepsilon)$ does not decrease.

Let us show that the probability of this event also does not decrease under the following transformation: let bets at all active vertices be equal to a (we still consider these vertices as active even if $a \notin M$). Denote by $p(k,m)$ the probability to collect at most k wins in m moves if the winning probability at each move is $(a+1)/2$. Clearly, it suffices to prove the following statement: for any k, the probability to collect at most k wins on the way to a leaf from a vertex v of active height h is not greater than $p(k, n-h)$ before the transformation and becomes equal to $p(k, n-h)$ after it. We prove it by inverse induction on the distance from v to the root. The induction step is obvious since the transformation does not increase the winning probability.

Since the number of wins in the obtained tree has a Bernoulli distribution with success probability $(a+1)/2$, we may use the large deviation probability bound. We obtain that the probability of $A(n,\varepsilon)$ is at most $\exp_e(-2n(\varepsilon/2)^2) = \exp_e(-0.5n\varepsilon^2)$. Summing over all n from $L\varepsilon$ to L, we get the desired bound on the probability of the first event in the statement of the lemma. The bound for the probability of the second event is proved similarly. Note that another way for proving such bounds is proposed in [1, Proposition 1]. Lemma 1 is proved.

The *capital $K(M)$ gained on M* is the capital that would be gained if only bets belonging to M were counted.

Lemma 2. *Let M, a, b, n, and d be the same as in Lemma 1. Then, for any $\varepsilon \in (0, 0.1]$, if $n \geq L\varepsilon$, $L > 20/\varepsilon^8$, and $a \geq \varepsilon$, the probability that $K(M) > db + 3L\varepsilon$ is not greater than $\exp_e(-0.3L\varepsilon^6)$.*

Proof. Divide the segment $[a,b]$ into the following parts: semiintervals $[a_1,b_1)$, $[a_2,b_2)$, $\ldots$, $[a_{m-1},b_{m-1})$ of length ε and a segment $[a_m,b_m]$ of length at most ε, where $a_1 = a$, $a_2 = b_1$, $\ldots$, $a_m = b_{m-1}$, $b_m = b$, $m \leq 1/\varepsilon$. By bets we mean only bets belonging to M. We exclude from consideration all parts that contain less than $L\varepsilon^2$ bets (note that the total number of bets on them is less than $L\varepsilon$,

and the total capital is not greater than the number of bets). We enumerate the remaining parts from left to right and consider the notation a_i, b_i, n_i, and d_i as related to the ith part.

By Lemma 1, with probability not less than $1-\exp_e(-0,4L\varepsilon^6)$, each d_i is nonnegative; hence, with probability not less than $1-\exp_e(-0,3L\varepsilon^6)$, all d_i are nonnegative. Considering the worst case for Mathematician (where all wins are due to the bet b_i, and all losses are due to the bet a_i), we conclude that, with the same probability, the capital gained on the ith part is not greater than $b_i(d_i+\frac{n_i-d_i}{2})-a_i\frac{n_i-d_i}{2}\le d_ib_i+n_i\varepsilon$. Summing over all i and recalling the excluded parts, we find that, with the above probability,

$$K(M)\le b\sum_i d_i+\varepsilon\sum_i n_i+L\varepsilon\le b(d+L\varepsilon)+L\varepsilon+L\varepsilon\le bd+3L\varepsilon.$$

Lemma 2 is proved.

Recall that the guaranteed capital on $[0,1]$ equals $L\delta\ln 2$. Since the selection rule is aimed at a large deviation of the fraction of zeros from $1/2$, we (bearing in mind the possibility to transform the strategy into a monotonic rule) split the strategy into two: one makes bets on $t_i=0$ only, and the other on $t_i=1$ only. At least one of these strategies guarantees the capital of $(L\delta\ln 2)/2$ on $[0,1]$. Denote this strategy by R.

Lemma 3. *Let Mathematician use strategy R,*

$$\alpha=\sqrt{0.5\ln 2(1-\delta^{0.05})\frac{\delta}{\ln(1/\delta)}}.$$

Then the probability to gain capital greater than $\dfrac{L\delta\ln 2}{2\ln(1/\delta)}$ on the segment $[0,\alpha]$ is not greater than $\exp_e(-L\delta^{6.7})$.

Proof. Put $K=\dfrac{L\delta\ln 2}{2\ln(1/\delta)}$. Clearly, on the segment $[0,\delta^{1.1}]$ it is possible to gain capital of at most $L\delta^{1.1}$; therefore, it remains to estimate the probability to gain capital greater than $K-L\delta^{1.1}$ on the segment $\Delta=[\delta^{1.1},\alpha]$, to which we refer our usual notation n and d. Since this capital can only be gained if $n>L\delta^{1.1}$, by Lemma 2 (with $\varepsilon=\delta^{1.1}$) we get that the probability that the gained capital is greater than $d\alpha+3L\delta^{1.1}$ does not exceed $\exp_e(-0.3L\delta^{6.6})$. Furthermore, by Lemma 1 (with $\varepsilon=\delta^{1.1}$), with probability not less than $1-\exp_e(-0.4L\delta^{3.3})$, we have $d\le n(\alpha+\delta^{1.1})$. Thus, with probability not less than $1-\exp_e(-L\delta^{6.7})$, the capital gained on Δ is at most $n(\alpha+\delta^{1.1})\alpha+3L\delta^{1.1}$. It is easily seen that this is less than $K-L\delta^{1.1}$ for small δ. Lemma 3 is proved.

This lemma allows us to concentrate upon bets that are not less than α. In particular, taking into account Lemma 1, Lemma 3 guarantees that our selection rules are efficient enough, at least from the point of view of the deviation of the fraction of zeros in the constructed subsequences from $1/2$: for length not less than $0.5L\delta\ln 2\Big(1-\dfrac{1}{\ln(1/\delta)}\Big)$, this deviation is with high probability not too much less than $(1+\alpha)/2$.

Divide the segment $[\alpha,1]$ into the following parts: semiintervals $[a_1,b_1)$, $[a_2,b_2)$, …, $[a_{m-1},b_{m-1})$ and a segment $[a_m,b_m]$, where $a_1=\alpha$, $a_2=b_1$, …, $a_m=b_{m-1}$, $b_m=1$, $b_i/a_i=k$ for all $i=1,\dots,m-1$, and $b_m/a_m\le k$,

$$k=\frac{1}{1-\beta}\left(1-\frac{1.2\ln\ln(1/\delta)}{\ln(1/\delta)}\right).$$

Note that $k>1$ for $\beta=\dfrac{2\ln\ln(1/\delta)}{\ln(1/\delta)}$. Let us find an upper bound on the number m of parts. We have $\alpha k^{m-1}\le 1/k$; therefore,

$$m\le 1+\frac{\ln(1/\alpha)}{\ln k}\le 1+\frac{0.5(\ln(2\operatorname{lb}e)-\ln(1-\delta^{0.05})+\ln\ln(1/\delta)+\ln(1/\delta))}{\ln k}\le\frac{0.6\ln(1/\delta)}{\ln k}.$$

As usual, we refer the notation n_i and d_i to the ith part. The capital gained on it is denoted by K_i. Any subset M of the set of parts defines a selection rule: a symbol of $\boldsymbol{t}$ is selected if and only if the strategy R makes a bet from a part belonging to M. By definition, the specific deficiency of this rule conditioned to S is

$$\frac{2\operatorname{lb} e}{L}\left(\frac{\sum d_i}{2\sum n_i}\right)^2\sum n_i = \frac{\operatorname{lb} e}{2L}\frac{\left(\sum d_i\right)^2}{\sum n_i},$$

where all sums are over $i \in M$. Thus, we have to prove that with high probability there exists M for which $\dfrac{\left(\sum d_i\right)^2}{\sum n_i}$ is not less than $Z = 2L\ln 2(1-\beta)\dfrac{\delta}{\ln(1/\delta)}$. Let us compose M only from parts where $n_i \ge L\delta^{1.1}$; since the total capital on the remaining parts is less than $Lm\delta^{1.1}$, we exclude them from consideration. Denote the set of the remaining parts by M^*. The following lemma gives a lower bound on the sum $\sum\limits_i \dfrac{d_i^2}{n_i}$, where i "runs over" *all* parts from M^*.

Lemma 4. *With probability not less than* $1-\exp_e(-L\delta^{6.8})$, *we have*

$$\sum_{i\in M^*}\frac{d_i^2}{n_i} > \frac{L\delta\ln 2}{2k}\left(1-\frac{2}{\ln(1/\delta)}\right).$$

Proof. By Lemma 1 (with $\varepsilon = \delta^{1.1}$), with probability not less than $1-\exp_e(-0.4L\delta^{3.3})$, for $i \in M^*$ we have the inequality $d_i \ge n_i(a_i - \delta^{1.1})$. By Lemma 2 (with $\varepsilon = \delta^{1.1}$), with probability not less than $1-\exp_e(-0.3L\delta^{6.6})$, we have $K_i \le d_i b_i + 3L\delta^{1.1}$, i.e., $d_i b_i \ge K_i - 3L\delta^{1.1}$. Hence, with probability not less than $1-\exp_e(-L\delta^{6.7})$, we have

$$\begin{aligned}\sum_{i\in M^*}\frac{d_i^2}{n_i} &\ge \sum_{i\in M^*} d_i\frac{n_i(a_i-\delta^{1.1})}{n_i} = \sum_{i\in M^*} d_i a_i - \delta^{1.1}\sum_{i\in M^*} d_i\\ &\ge \sum_{i\in M^*}\frac{d_i b_i}{k} - L\delta^{1.1} \ge \frac{1}{k}\left(\sum_{i\in M^*} K_i - 3Lm\delta^{1.1}\right) - L\delta^{1.1}.\end{aligned}$$

According to Lemma 3 and to the definition of M^*, with probability not less than $1-\exp_e(-L\delta^{6.7})$, we have

$$\sum_{i\in M^*} K_i \ge \frac{L\delta\ln 2}{2}\left(1-\frac{1}{\ln(1/\delta)}\right) - Lm\delta^{1.1}.$$

Let us lower bound $\ln k$, taking into account the relation between β and δ. We have

$$\ln k = \ln\frac{1}{1-\beta} + \ln\left(1-\frac{1.2\ln\ln(1/\delta)}{\ln(1/\delta)}\right) \ge \beta - \frac{1.3\ln\ln(1/\delta)}{\ln(1/\delta)} \ge \frac{0.7\ln\ln(1/\delta)}{\ln(1/\delta)}.$$

Hence, with probability not less than $1-\exp_e(-L\delta^{6.8})$, for small δ we have

$$\begin{aligned}\sum_{i\in M^*}\frac{d_i^2}{n_i} &\ge \frac{L\delta\ln 2}{2k}\left(1-\frac{1}{\ln(1/\delta)}\right) - \frac{Lm\delta^{1.1}}{k} - \frac{3Lm\delta^{1.1}}{k} - L\delta^{1.1}\\ &\ge \frac{L\delta\ln 2}{2k}\left(1-\frac{1}{\ln(1/\delta)}\right) - \frac{4L\delta^{1.1}\ln^2(1/\delta)}{\ln\ln(1/\delta)} - L\delta^{1.1} > \frac{L\delta\ln 2}{2k}\left(1-\frac{2}{\ln(1/\delta)}\right).\end{aligned}$$

Lemma 4 is proved.

Lemma 5. *With probability not less than* $1-\exp_e(-L\delta^{6.8})$ *there exists a subset* M *of* M^* *such that*

$$\frac{\left(\sum\limits_{i\in M} d_i\right)^2}{\sum\limits_{i\in M} n_i} \ge Z.$$

Proof. By Lemma 4, with probability not less than $1 - \exp_e(-L\delta^{6.8})$, we have

$$\sum_{i\in M^*} \frac{d_i^2}{n_i} > \frac{L\delta \ln 2}{2k}\left(1 - \frac{2}{\ln(1/\delta)}\right).$$

Assume that a required M does not exist. Then for any $J \subseteq M^*$ we have

$$\left(\sum_{i\in J} \frac{d_i}{n_i} n_i\right)^2 \le Z\left(\sum_{i\in J} n_i\right).$$

Repeating the arguments from the proof of Proposition 2 in [1], we obtain the inequality

$$\sum_{i\in M^*} \frac{d_i^2}{n_i} < Z + \frac{Z}{4}\left(\ln\left(\sum_{i\in M^*} n_i\right) - \ln Z\right).$$

The proof repeats that of [1] word-by-word; the only slight difference is in justification of the inequality $\sum\limits_{i\in M^*} n_i > Z$, which in our case looks as follows:

$$\sum n_i = \sum \frac{n_i^2}{n_i} \ge \sum \frac{d_i^2}{n_i} > \frac{L\delta \ln 2}{2k}\left(1 - \frac{2}{\ln(1/\delta)}\right) > Z.$$

Hence,

$$\frac{L\delta \ln 2}{2k}\left(1 - \frac{2}{\ln(1/\delta)}\right) < (1-\beta)\frac{2L\delta \ln 2}{\ln(1/\delta)}\left(1 + \frac{1}{4}\ln\frac{L}{Z}\right),$$

i.e.,

$$\frac{1}{k}\left(1 - \frac{2}{\ln(1/\delta)}\right) < (1-\beta)\frac{4 + \ln\ln(1/\delta) - \ln 2 - \ln\ln 2 + \ln(1/\delta) - \ln(1-\beta)}{\ln(1/\delta)},$$

which is wrong if

$$k \le \frac{\ln(1/\delta) - 2}{(1-\beta)(4 + \ln\ln(1/\delta) - \ln\ln 4 + \ln(1/\delta) - \ln(1-\beta))},$$

in particular, if

$$k(1-\beta) \le \frac{\ln(1/\delta) - 2}{\ln(1/\delta) + 1.1\ln\ln(1/\delta)}.$$

It is easy to verify by direct substitution that, for our β and k (if $\delta \le \exp_e(-e^{50})$), this inequality is satisfied. Lemma 5 is proved.

The proof of Theorem 2 follows the same scheme as in [1]. Namely, the small subset of S (called *exceptional*) on which none of the described rules acts is declared to be the new set S, and the whole construction is iterated on it. There arises a smaller exceptional set, and so on. It remains to estimate the number of rules. Using the previously obtained estimate for $\ln k$, we have

$$\operatorname{lb} k = \operatorname{lb} e \ln k \ge \operatorname{lb} e\left(\beta - \frac{1.3\ln\ln(1/\delta)}{\ln(1/\delta)}\right) = \frac{\beta\ln(1/\delta) - 1.3\ln\ln(1/\delta)}{\ln 2\ln(1/\delta)}.$$

Therefore, the number of rules at each iteration is not greater than

$$2^m \le \exp_2\left(\frac{0.6\operatorname{lb}(1/\delta)}{\operatorname{lb} k}\right) = \left(\frac{1}{\delta}\right)^{\frac{0.6\ln 2\ln(1/\delta)}{\beta\ln(1/\delta) - 1.3\ln\ln(1/\delta)}}.$$

It follows from Lemma 5 that the number of iterations is at most $(1/\delta)^{6.8}$. Theorem 2 is proved.

Of special interest might be selection rules for which the set of accounted bets is connected and simply described. Therefore, of interest might be the question of corresponding estimates on the specific deficiency and the number of rules for which this set consists of one segment (let us call them *segment rules conditional to* S). The result of An.A. Muchnik formulated in Theorem 1 in fact states that, for δ' of the order of δ^2, for a δ'-covering of a set S (where $\delta(S)=\delta$), a linear (in $1/\delta$) number of monotonic segment rules conditional to S (more precisely, threshold rules, i.e., those with the segment of accounted bets of the form $[p,1]$) is sufficient.

Our technique allows to obtain an estimate for segment rules that cover though not all the set S but almost all. More precisely, we have the following result.

Theorem 3. *Let a number $\delta\in(0,\exp_e(-e^{50}))$ and a positive integer $L\ge(1/\delta)^5$ be fixed. Consider sets of binary sequences of length L. For an arbitrary set S of specific deficiency not less than δ there exists a family of at most $0.6\ln(1/\delta)$ monotonic segment rules conditional to S that δ'-cover almost all S except for, possibly, a subset of cardinality not greater than $|S|\exp_e(-L\delta^{6.8})$, where $\delta'=\dfrac{1}{2.5e}\dfrac{\delta}{\ln(1/\delta)}$.*

Proof. We construct the set of parts in the same way as in the proof of Theorem 2 but with $k=e$. Let us only consider rules for which the set of bets is some segment of the form $[a_i,b_i]$. It follows from Lemma 3 that, with probability not less than $1-\exp_e(-L\delta^{6.7})$, on some segment there is gained capital $K_i\ge\dfrac{L\delta\ln 2}{2m}\left(1-\dfrac{1}{\ln(1/\delta)}\right)$. The rule corresponding to it has specific deficiency $D=\dfrac{\operatorname{lb}e}{2L}\dfrac{d_i^2}{n_i}$. By Lemma 1 (with $\varepsilon=\delta^{1.1}$), with probability not less than $1-\exp_e(-0.4L\delta^{3.3})$, we have $d_i\ge n_i(a_i-\delta^{1.1})$. Hence, with the same probability, $D\ge\dfrac{\operatorname{lb}e}{2L}n_i(a_i-\delta^{1.1})^2$. On the other hand, by Lemma 2 (with $\varepsilon=\delta^{1.1}$), with probability not less than $1-\exp_e(-0.3L\delta^{6.6})$, we have $K_i\le d_ib_i+3L\delta^{1.1}$, i.e., $d_i\ge\dfrac{K_i-3L\delta^{1.1}}{b_i}$. Hence, with the same probability, $D\ge\dfrac{\operatorname{lb}e}{2L}\dfrac{(K_i-3L\delta^{1.1})^2}{b_i^2n_i}$. Since the first estimate for D increases with n_i, and the second decreases, the smallest possible value of D is attained if $n_i(a_i-\delta^{1.1})^2=\dfrac{(K_i-3L\delta^{1.1})^2}{b_i^2n_i}$, i.e., if $n_i=\dfrac{K_i-3L\delta^{1.1}}{b_i(a_i-\delta^{1.1})}$. For this n_i, taking into account estimates for K_i and m and using the inequalities $b_i/a_i\le k$, $b_i>\alpha$, we obtain that, with probability not less than $1-\exp_e(-L\delta^{6.8})$, for small δ we have

$$D\ge\frac{\operatorname{lb}e}{2L}\frac{(a_i-\delta^{1.1})(K_i-3L\delta^{1.1})}{b_i}\ge\frac{\operatorname{lb}e(a_i-\delta^{1.1})}{2Lb_i}\frac{L\delta\ln 2}{1.2\ln(1/\delta)}\left(1-\frac{2}{\ln(1/\delta)}\right)$$
$$\ge\frac{(1/e)-\sqrt{\delta}}{2.4}\frac{\delta}{\ln(1/\delta)}\left(1-\frac{2}{\ln(1/\delta)}\right)\ge\frac{1}{2.5e}\frac{\delta}{\ln(1/\delta)}.$$

The number of rules is at most $m\le 0.6\ln(1/\delta)$. Theorem 3 is proved.

Remark. It is easily seen that, by reducing the range of possible values of δ, the constant 0.6 in the statement of Theorem 2 can be made arbitrarily close to 0.5; the constant 6.8 in Theorems 2 and 3, arbitrarily close to 6; and the constant 2.5 in Theorem 3, to 2.

Let us now present a proof of Theorem 1.

Proof of Theorem 1. As was already mentioned, all our rules will be threshold rules; i.e., decision on selecting a current bit to a subsequence is made if the bet (computed according to the rule described in the beginning of the proof of Theorem 2) is not less than a certain threshold. Let us say that an ordered pair $\langle r_1,r_2\rangle$ of two integers is good if $r_2-r_1\le\varepsilon\delta$, where ε is the number that determines the relation between c_1 and c_2 (for example, $\varepsilon=0.1$). As we know, there exists a strategy R which, first, matches the selection rule (in the sense that it "knows" what bits it should select more) and, second, gains capital of at least $0.5L\delta\ln 2$. Thus, if we do not count some set of good pairs $\langle r_1,r_2\rangle$ of bets, where the bet r_1 losses, then the counted capital is still at least $L\delta(0.5\ln 2-\varepsilon)$.

Assume that a game played according to strategy R is over and all bets are marked on the real line, with indication of wins and losses. Let us delete pairs of bets according to the following rule: look through bets in the ascending order and, for each (not yet deleted) losing bet r_1, check whether there exists a (not yet deleted) winning bet r_2 forming a good pair $\langle r_1, r_2\rangle$; if exists, take the largest of such bets r_2 and delete the pair $\langle r_1, r_2\rangle$. It is clear that, after all deletions, all remaining losses are less than all remaining wins, and the interval between these two sets (denote them by P and V, respectively) is greater than $\varepsilon\delta$. Let us set thresholds beginning from 0, with step $\varepsilon\delta$; then the number of rules is approximately $(\varepsilon\delta)^{-1}$. If P is empty, consider the threshold $T = 0$; otherwise, consider the smallest threshold T in the interval between P and V. Let us show that, for any deleted pair $\langle r_1, r_2\rangle$, we cannot have $r_2 < T \le r_1$. Indeed, if we assume this possibility, then, according to the way of deleting pairs, the winning bet r_2 would have already been paired with the largest element of P before the deletion, and this element would have been deleted.

Thus, in any case, the difference d between the number of actual (i.e., lying not below the threshold T) wins and losses is not less than $|V|$. Taking into account that $|V| \ge L\delta(0.5\ln 2 - \varepsilon)$, we get the following bound on the specific deficiency D of the rule (here n is the length of the selected subsequence):

$$D = \frac{2\,\mathrm{lb}\,e}{L}\left(\frac{d}{2n}\right)^2 n = \frac{d^2\,\mathrm{lb}\,e}{2Ln} \ge \frac{L^2\delta^2(0.5\ln 2 - \varepsilon)^2}{2L^2} = 0.5\,\mathrm{lb}\,e(0.5\ln 2 - \varepsilon)^2\delta^2,$$

which proves the theorem.

The author is deeply grateful to An.A. Muchnik, who attracted his attention to this subject and made many valuable remarks, which helped the author to considerably improve the text.

REFERENCES

1. Muchnik, An.A. and Semenov, A.L., On the Role of the Law of Large Numbers in the Theory of Randomness, *Probl. Peredachi Inf.*, 2003, vol. 39, no. 1, pp. 134–165 [*Probl. Inf. Trans.* (Engl. Transl.), 2003, vol. 39, no. 1, pp. 119–147].

2. Shiryaev, A.N., *Veroyatnost'*, vol. 1: *Elementarnaya teoriya veroyatnostei. Matematicheskie osnovaniya. Predel'nye teoremy* (Elementary Probability Theory. Mathematical Foundations. Limit Theorems), Moscow: MCCME, 2004, 3rd ed. Second edition translated under the title *Probability*, New York: Springer, 1996.

3. Uspensky, V.A., Semenov, A.L., and Shen', A.Kh., Can an Individual Sequence of Zeros and Ones Be Random?, *Uspekhi Mat. Nauk*, 1990, vol. 45, no. 1, pp. 105–162 [*Russian Math. Surveys* (Engl. Transl.), 1990, vol. 45, no. 1, pp. 121–189].

4. Muchnik, An.A., Semenov, A.L., and Uspensky, V.A., Mathematical Metaphysics of Randomness, *Theor. Comput. Sci.*, 1998, vol. 207, no. 1–2, pp. 263–317.