*Информационные процессы, Том 9, № 4, 2009, стр. 225–248.*
*© 2009 Sagalovich, Chilingarjan.*

**INFORMATION PROCESSES**

# Separating Systems and New Scopes of Its Application

**Yu.L. Sagalovich and A.G. Chilingarjan**

*Institute for Information Transmission Problems, Russian Academy of Sciences, Moscow, Russia*

Received 14.10.2009

**Аннотация**—This article is motivated by the fact that, at different times, one and the same theory of separating systems has served as an adequate research technique for such different areas of science and technology as automata synthesis, technical diagnosis, and the construction of hash functions. The genetics problems are maybe the new field of its application

## 1. INTRODUCTION

The study of separating systems was apparently originated in the problems of critical race-free coding for states of discrete automata. These problems are connected with the so-called races of memory cells of automata, i.e., with the fact that, during the transition from one stable inner state, $s_1$, to another, $s_2$, the binary memory cells of an asynchronous automaton do not change their states simultaneously. This transition will be denoted as $s_1 \to s_2$. The collection of states of memory cells of the automaton at a given time instant forms its inner state.

A consequence of this time nonuniformity in the course of the transition from one state to another is the rise of intermediate states $s_{12}^{(a)}, s_{12}^{(b)}, \ldots, s_{12}^{(v)}$. The appearance of one and the same intermediate state $s_{12}^{(i)} = s_{12}^{(j)}$ in the course of two transitions $s_1 \to s_2$ and $s_3 \to s_4$ under one and the same input signal would violate the determinism principle. The designer would be stymied, not knowing whether the automaton should be "directed" from this common intermediate state to $s_2$ or to $s_4$. This situation is called the critical race. Researchers have put much effort into avoiding this situation by using an appropriate encoding of the automata states. The problem is, given two transitions $s_1 \to s_2$ and $s_3 \to s_4$ that are possible for one and the same input signal, to find at least one bit in the binary encoding of $s_1, s_2, s_3, s_4$ that equals 0 for $s_1$ and $s_2$ and 1 for $s_3$ and $s_4$, or vice versa. Then for all the intermediate states $s'_{12}$ this bit would equal 0 and for all states $s'_{34}$ it would equal 1 (or vice versa). Thus, the transitions

$$s_1 \to s_2 \,\text{and}\, s_3 \to s_4 \tag{1}$$

would be *separated*, which gave rise to the title "separating systems." It should be noted that races (and, moreover, critical races) will never occur if the transitions (1) involve the state change of only one memory cell, which yields no intermediate states. In the extensive Russian literature on this subject, the encoding of automata states in order to avoid critical races is often called the "displacement of states."

The numerous references on the encoding that involves, in one way or another, a transition table or another means of defining the automaton are surveyed in [12, 45]. Here we fix our attention on [25]. Its author was the first to consider a "universal" encoding (without paying attention to the specific features of the automaton), which assumes that the transitions (1) are separated irrespective of the input signal and irrespective of whether they take place at all. The separating system was formed by a binary equidistant MacDonald code [40] with the parameters $n = 2^m - 1, k = m, d = 2^{m-1}$, where $n$ is the length, $k$ the number of information symbols, and $d$ the minimum code distance. If $M$ is the size of the code, then $k = \log_2 M$.

Intensive research toward increasing automata reliability began nearly at the same time. It involved error-correcting codes [1, 11, 70]. This revealed a connection between the critical race-free coding and error-correcting coding. Indeed, if states $s_i$ lie at a distance $d$ apart, it is impossible that only one memory cell changes its state in the course of the transition $s_1 \to s_2$; therefore, intermediate states during transitions are inherently inevitable, and we must separate the transitions.

This led to the joint study of fault-tolerant and critical race-free encoding of automata states, which was initially stated in [46] in 1965. In this paper, a necessary and sufficient condition was given for the automaton to sustain the failure of any $t$ or less memory cells under their races (critical races).

Any ordered quadruple of vectors $s_1, s_2, s_3$, and $s_4$ must contain not less than $\theta \geq 2t+1$ "separating" coordinates. In other words, for an automaton to be tolerant toward the failure of $t$ or less memory cells it is necessary and sufficient that any ordered quadruple of vectors

$$\begin{aligned} s_1 &= (a_1^{(1)}, a_2^{(1)}, \ldots, a_i^{(1)}, \ldots, a_n^{(1)}),\\ s_2 &= (a_1^{(2)}, a_2^{(2)}, \ldots, a_i^{(2)}, \ldots, a_n^{(2)}),\\ s_3 &= (a_1^{(3)}, a_2^{(3)}, \ldots, a_i^{(3)}, \ldots, a_n^{(3)}),\\ s_4 &= (a_1^{(4)}, a_2^{(4)}, \ldots, a_i^{(4)}, \ldots, a_n^{(4)}), \end{aligned} \tag{2}$$

contain not less than $\theta \geq 2t+1$ so-called regular columns [34, 35] of the form

$$(0011)^T \qquad \text{or} \qquad (1100)^T.$$

Thus, the problem of constructing $(2, 2, 2t)$-separating systems (s.s.) was formulated. The symbol $2, 2$ denotes the separation of the two pairs of transitions and the symbol $2t$ recalls the existence of $2t + 1$ regular columns (separating coordinates). A first solution to this problem was suggested in the class of Hadamard matrices and various combinations of them in the above-mentioned paper [46]. This paper also contains relations between the six pairwise distances in the vector quadruple $s_1, s_2, s_3, s_4$ for various properties of the pair of transitions (1). Already there we noticed that the maximum distance $D$ acts on a level with the minimum distance $d$, which plays the key role in error-correcting codes. We also revealed a connection between the three quantities $\theta, d$, and $D$. Specifically, for some of the properties mentioned, including the case in which (1) contains only one transition, $s_1 \to s_2$ say, and $s_3 = s_4$, we have proved that $\theta$ satisfies the following inequality:

$$2\theta \geq 2d - D. \tag{3}$$

In fact, this side result initiated the study of $(2, 1, 2t)$-s.s., discussed in detail in Section 5. This symbol denotes that intermediate states can appear as system vectors under no unique transition.

We stress that the quantities $d$ and $\theta$ are different both in their numerical value and meaning.

In [28], the so-called $(2, 2)$ completely separating systems (c.s.s.) was introduced. C.s.s. differ from s.s. by the property that the quadruple of vectors (2) contains at least one column of the form $(0011)^T$ and at least one column $(1100)^T$. In addition to separating transitions, this property provides the monotonicity [64] and inversion-free implementation of automata functions. For an automaton to sustain the failure of any $t$ or less memory cells it is necessary that the quadruple contain at least $\theta$ columns of both forms. This gives rise to $(2, 2, 2t)$-c.s.s.

A.D. Fridman and others [10], with a reference to [35], generalized the concept of $(2, 2)$-s.s. $(t \geq 0)$ to a general notion of $(i, j)$-s.s. In our notation, this is a set $Q$ of $M$ binary vectors such that for any two of its disjoint subsets $Q_1$ and $Q_2$ of size $i$ and $j$, respectively, there exists a digit (a coordinate), in which all of the vectors in $Q_1$ contain a symbol opposite to the one contained in all of the vectors in $Q_2$.

In fact, one could think of going even further and requiring, as in [47, 48], that in the mentioned digit (coordinate) the vectors of $Q_1$ (respectively, $Q_2$) contain some given combinations of symbols. The coding problem for automata states gave rise to $(3,3)$-s.s. (see [66]).

## 2. EXISTENCE BOUNDS

A first lower bound for $(2,2)$-s.s. was obtained in [10] in 1969 for the case of $t = 0$, i.e., $\theta \geq 1$). This bound states that asymptotically over $n$ there exists a $(2,2)$-s.s. with $M$ vectors of length $n$ provided that

$$R = (1/n)\log_2 M < .0481. \tag{4}$$

A first asymptotic existence bound, which relates the two quantities $\theta/n$ and $R = (1/n)\log_2 M$ as $n \to \infty$, was obtained in [38, 39]. As a side result, these papers suggested that one consider $q$-ary memory cells instead of binary ones without thinking of their actual existence. This, in turn, led to various hypothetical models for races (critical races) and to the corresponding existence bounds for linear and nonlinear $(2,2,2t)$-s.s. [39, 48]. In this survey, we shall not describe the race models for multivalued memory cells. We only mention briefly that memory cells take on values in $GF(q)$ and, apart from the difference in the times of state changes, they can pass, or not pass, through intermediate states. If they do, this transition can follow, or not follow, a certain pattern. For simplicity, let us fix three models. In Model 1, we assume that the cells change their states at different time instants and in the course of this change each of them can take on any value in $GF(q)$. Model 2 is characterized only by the time nonuniformity while intermediate values are forbidden. Finally, Model 3 suggests that the memory cell can take on the intermediate values, though only in a given order. There also exist some other models, which we do not describe here.

To get rid of the critical races under a possible failure of any $t$ or less memory cells, it is necessary and sufficient that the ordered quadruple of vectors (2) contain not less than $\theta$ columns that satisfy the following conditions:

for Model 1,

$$a_i^{(1)} = a_i^{(2)}; \quad a_i^{(3)} = a_i(4); \quad a_i^{(1)} \neq a_i^{(3)}; \tag{5}$$

for Model 2,

$$a_j^{(j)} \neq a_i^{(m)};\ j = 1,2;\ m = 3,4; \tag{6}$$

for Model 3,

$$\begin{array}{l} a_i^{(j)} \neq a_i^{(m)};\ j = 1,2;\ m = 3,4, \text{ and the intervals between } a_i^{(1)}, a_i^{(2)} \\ \text{and } a_i^{(3)}, a_i^{(4)} \text{ have no points in common.} \end{array} \tag{7}$$

It is easy to see that for $q = 2$, all three models are identical, as are the stated necessary and sufficient conditions.

The columns of type (5), (6), and (7) are also "separating." In [47, 48], they were also called "regular" while the remaining ones were called "irregular." It is easy to calculate that the number of regular columns among all possible $q^4$ columns equals

$$\alpha = \begin{cases} q(q-1) \quad \text{for Model 1} \\ q(q-1)^2 + q(q-1+(q-2)^2 \quad \text{for Model 2} \\ q^2(q^2-1)/6 \quad \text{for Model 3.} \end{cases} \tag{8}$$

Accordingly, the fraction $\beta = \alpha/q^4$ of regular columns among all $q^4$ columns equals

$$\beta = \begin{cases} (q-1)/q^3 \\ 1-4/q+6/q^2-3/q^3, 1/6-1/6q^2. \end{cases} \tag{9}$$

This fraction plays an essential role in the behavior of existence bounds.

Let $R = (1/n)\log_q M$. In [38, 39], we have used random selection and counting argument in order to obtain the following (asymptotic over $n$) lower bound, i.e., an existence bound for the $(2,2,2t)$-s.s.:

$$R < -(1/3)\{(\theta/n)\log_q \beta + (1-(\theta/n))\log_q(1-\beta) + H(\theta/n)\}, \tag{10}$$

where $H(x) = -x\log_q x - (1-x)\log_q(1-x)$.

It is noteworthy that initially, in [38, 39], the bound (10) was proved for linear systems for $q = 2$ and $q = 3$, while for nonlinear $(2,2,2t)$-s.s. we had 1/3 in place of 1/4. This was due to the fact that initially for the nonlinear case we took into account all the $\binom{M}{4}$ quadruples of vectors (2), while in the linear case we could manage with just $\binom{M}{3}$ triples

$$0,\ s_2,\ s_3,\ s_4, \tag{11}$$

since in this case, $s_1$ can always be assumed to be zero. The bound remained in this "double" form till 1983, when we published Pinsker's proof in the Appendix to [57]. This proof enabled one to take into account in the nonlinear case only the triples of vectors. This proof is important especially for Models 1 and 3, since for them linear $(2,2,2t)$-s.s. do not exist, and the only available bound had been the one with 1/4 instead of 1/3.

Equation (10) implies that $R$ is a positive constant whenever $\theta/n \le \beta$ for all three race models. Notice that the quantity $\theta/n$ recalls the relative code distance for usual block error-correcting codes. Thus, it can be called the relative "generalized" code distance. Note also that the usual block code correcting $t$ independent errors can be called a $(1,1,2t)$-separating system.

The curve that bounds the domain (10) intersects the axis $\theta/n$ at the point $(\beta, 0)$ and the axis $R$ at the point

$$(0, -(1/3)\log_q(1-\beta)). \tag{12}$$

From Eq. (8) we easily deduce that for $q = 2$, we have $\beta = 1/8$ for all three models.

The second extreme point for $q = 2$ can be found by substituting into (12) the value $\beta = 1/8$. This gives $R = (3 - \log 7)/3 = .0642$. This is greater than (4) by a factor of 4/3. The reason for this lies in the mentioned difference in derivations for linear and nonlinear systems. In [10], the authors did not make use of the linearity and were not aware of the method of reducing quadruples (2) to triples (11) for nonlinear systems. Consequently, instead of 1/3 in (12) they used 1/4.

In the above discussion, we have emphasized that the system linearity and related consideration of triples instead of quadruples yields an improvement to the bound only for $q = 2$ and $q = 3$. The reason for this is that the proof of lower bounds for linear systems reveals certain specific features of the linear dependence over $GF(q)$ of the vectors in the triple

$$s_2, s_3, s_4. \tag{13}$$

One has to distinguish among the following four cases (up to the order of vectors) of linear dependence of vectors in (13):

1. $s_2 = \xi s_3 = \zeta s_4, \quad , \xi, \zeta \neq 1$;
2. $s_2 = \xi s_3$ (or $s_2 = \xi s_4$), $\quad \xi, \zeta \neq 1$;
3. $s_3 = \xi s_4, \quad \xi \neq 1$;
4. $s_2 = \xi s_3 + \zeta s_4$ and either (a) $\xi = \zeta = 1$, or (b) $\xi, \zeta, \xi + \zeta \neq 1$, or (c) $\xi = 1, \zeta \neq 1$ or $\xi \neq 1, \zeta = 1$, or (d) $\xi, \zeta \neq 1, \xi + \zeta = 1$.

Naturally, everywhere $\xi, \zeta \neq 0$. We observe that for $q = 2$ only case 4(a) is possible; for $q = 3$ the possibilities are cases 2,3,4(a), 4(c), and 4(d). Cases 1 and 4(b) are possible only for $q > 3$.

The number $\alpha$ of regular columns for all these cases of linear dependence equals

$$\alpha = \begin{cases} (q-1) \quad \text{for case 1} \\ (q-1)(q-2) \quad \text{for cases 2,3,4(c)), and 4(d)} \\ (q-1)^2 \quad \text{for case 4(a)} \\ (q-1)(q-3) \quad \text{for case 4(b).} \end{cases}$$

The corresponding values of $\beta$ are

$$\beta = \begin{cases} (q-1)/q \quad \text{for case 1} \\ (q-1)(q-2)/q^2 \quad \text{for cases 2,3,4(c), and 4(d)} \\ (q-1)^2/q^2 \quad \text{for case 4(a)} \\ (q-1)(q-3)/q^2 \quad \text{for case 4(b).} \end{cases}$$

Considering all these cases of linear dependence and the fraction $\beta$ of pairs or single vectors that exhibit it we see that the linear dependence does not affect the existence bound for $q = 2$ and $q = 3$. For $q = 4$, the lower bound for linear systems is inferior to that for nonlinear ones. For $q = 5$, the bound is formed by two branches given by the intersection of the bounds for the linear and nonlinear cases at the point $\theta/n \approx .29998$. For $q \to \infty$, the bounds coincide and are given by the equality

$$R = (1/3)(1 - \theta/n). \tag{14}$$

The existence bound is usually called lower. This formally contradicts the sign $<$ in (10) and further equations. One should use the sign $\geq$. However, we prefer to employ the notation suggested above and to accompany the inequalities with the wording "... exists if the following inequality holds ...<... ."

For the binary $(2, 2, 2t)$-c.s.s., we have obtained the following existence bound [57]:

$$R < (1/3)(4 - H(\theta/n) - (1 - \theta/n)\log_2 15). \tag{15}$$

In [4], it is shown that linear c.s.s. do not exist.

In the binary case, the problem of constructing nonlinear $(2, 2, 2t)$-s.s. was formulated anew in [41] in 1973. This paper also gave the existence bound. The only difference of this bound from (10) was that it was obtained only for $q = 2$ and with a weaker factor 1/4 in place of 1/3.

The authors of [17, 22, 33] considered the so-called "double" $(2, 2, 2t)$-s.s. Their idea is as follows. The set $s$ of automata states splits into two (probably intersecting) subsets $S_1$ and $S_2$ of size $M_1$ and $M_2$ . The only possible transitions are those from the state $s^1 \in S_1$ to $s^2 \in S_2$ or vice versa. To every state in $S_i$ $(i = 1, 2)$ corresponds a binary vector $y$ of length $n$. Thus, we assign the sets of vectors $Y^1$ and $Y^2$ to the sets $S_1$ and $S_2$. Since the subsets $S_1$ and $S_2$ are not necessarily disjoint, some states may be assigned with two binary vectors. However, this does not prevent us from maintaining transitions and following the encoding principle. Hence, one does not need to separate the pairs of vectors within both sets $Y^1$ and $Y^2$. Thus, it is sufficient to separate only those pairs of vectors $y$ that are contained in different sets $Y^1$ and $Y^2$. The set pair $Y^1, Y^2$ was called the double $(2, 2, 2t)$-s.s. in [10]. The minimization problem of $n$ for given $M_1, M_2$, and $\theta$ is equivalent to the maximization problem of $M_2$ for given $M_1, n$, and $\theta$.

A straightforward generalization in [10] led to an existence bound for $(i, j)$-s.s. of the form $k/m < (i + j)/\log_2(1/(1 - 2^{1-i-j}))$.

## 3. UPPER BOUNDS

The upper bounds [47, 48] are based on a number of theorems. Before stating them, we give the following definition.

Suppose the vectors of a code $A$ of length $n$ are written as the rows of an $M$ by $n$ matrix. Taking a subset of columns of this matrix and all or some rows in this subset, we get another code. Let us denote this code by $B$ and say that $A$ contains $B$ or $B$ is contained in $A$.

**Theorem 1.** *For Models* 1 *and* 3, *any* $(2, 2, 2t)$*-s.s. contains a usual error-correcting code of length* $d$ *equal to the minimum distance of this* $(2, 2, 2t)$*-s.s., size not less than* $M - 2$ *and minimum distance* $\delta$ *not less than* $4t + 2$.

**Theorem 2.** *For Model* 2, *any* $(2, 2, 2t)$*-s.s. contains a usual error-correcting code of length* $d$ *equal to the minimum code distance of this* $(2, 2, 2t)$*-s.s., size not less than* $M$ *(respectively,* $M - 2$*) and distance* $\delta$ *not less than* $2t + 1$ *(respectively,* $4t + 2$*) for* $q > 3$ $(q = 2, 3)$.

In addition, it is proved that if the $(2, 2, 2t)$-s.s. is linear, the code that it contains, according to Theorem 2, is also linear. We call this code an $(M, d, \delta)$ code.

**Theorem 3.** *For Model* 3, *the linear* $(M, d, \delta)$ *code contained in a linear* $(2, 2, 2t)$*-s.s., according to Theorem 2, itself contains a linear code of length* $\delta$ *equal to the minimum distance of the* $(M, d, \delta)$ *code, size* $M$ $(M/2)$ *for* $q > 2$ $(q = 2)$, *and minimum distance not less than* $2t + 1$.

Thus, one and the same quantity $d$ is at the same time the minimum distance of a $(2, 2, 2t)$-s.s. and the length of a code contained in this system. In the same way, $\delta$ equals the code distance of a code of length $d$ and the length of another code contained in the former one. This fact serves as the basis for obtaining upper bounds.

The principal feature of the proofs of the theorems is that $s_1$ is taken to be all-zero (if this vector is not present, it can always be formed), and $s_3$ is a vector of minimum weight $d$. The vectors $s_2$ and $s_4$ are chosen from the remaining $M - 2$ vectors in an arbitrary manner. Later these two vectors are interchanged, which yields two quadruples of vectors $s_1, s_2, s_3, s_4$ and $s_1, s_4, s_3, s_2$. If the original system is linear, jointly with the vectors $s_1$ and $s_3$ we choose the vector $s_4$, which has minimal weight in the coordinates corresponding to the nonzeros of $s_3$. Details of the proofs are found in [47, 48].

According to Theorems 1–3, we have $d = \lambda n, \delta = \mu d = \mu\lambda n$, where for $M$ sufficiently large, $0 \le \lambda, \mu \le (q - 1)/q$ (see [47, 48]).

Put $k = \log_q M$ and use the representation of upper bounds for the usual error-correcting codes in the form $R \le f(d/n)$ to obtain the relation

$$k/n \le f(\lambda). \tag{16}$$

In addition, we use Theorems 1–3 to obtain the following relations.

For nonlinear $(2, 2, 2t)$-systems for Models 1 and 3 and all $q$, and for linear $(2, 2, 2t)$-s.s. for Model 2 and $q = 2$ or $q = 3$, we have

$$k/n \leq \lambda f(4t/\lambda n). \tag{17}$$

Moreover, the equality $\delta = \mu d$ implies

$$k/n \leq \lambda f(\delta/\lambda n) = \lambda f(\mu). \tag{18}$$

Consider arbitrary (linear or nonlinear) $(2, 2, 2t)$-s.s. in the case of Model 2. Let $q > 3$. Then $k/d \leq f(2t/d)$ or

$$k/n \leq \lambda f(2t/\lambda n). \tag{19}$$

In addition, Eq. (18) also holds true.

Consider linear $(2, 2, 2t)$-s.s. in the case of Model 2 and arbitrary $q$. Then $k/\delta \leq 2t/\delta$ or

$$\mu k/n \leq /\mu\lambda(2t/\mu\lambda n). \tag{20}$$

(The fact that for $q = 2$, the size of a code of length $\delta$ is two times less than the size of a $(2, 2, 2t)$-s.s. (cf. Theorem 3) does not affect the asymptotic behavior of $R$ since it decreases $R$ by the quantity $1/n$. The same reason applies to the substitution of $M$ for $M - 2$ (cf. Theorems 1 and 2) since $M$ is large.)

Let us summarize the above argument.

3.1. Consider nonlinear $(2, 2, 2t)$-s.s. For Models 1 and 3 and all $q$ and for Model 2 and $q = 2$ or $q = 3$ we have the following simultaneous inequalities:

$$k/n \leq f(\lambda); \quad k/n \leq \lambda f(2\theta/\lambda n). \tag{21}$$

3.2. Consider linear $(2, 2, 2t)$-s.s. For Model 2 and all $q$ we have the following simultaneous inequalities:

$$k/n \leq f(\lambda); \quad k/n \leq \lambda f(\theta/\lambda n). \tag{22}$$

3.3. Consider linear $(2, 2, 2t)$-s.s. For Model 2 and arbitrary $q$, we have the following simultaneous inequalities:

$$k/n \leq f(\lambda); \quad k/n \leq \lambda f(\mu); \quad k/n \leq \lambda\mu f(\theta/\lambda\mu n). \tag{23}$$

It is easy to see that for all three cases covered by the theorems, the quantities $\lambda(\theta/n)$ and $\mu(\theta/n)$, respectively, obey the following equations:

$$f(\lambda) = \lambda f(2\theta/\lambda n), \tag{24}$$

$$(\lambda) = \lambda f(\theta/\lambda n) \tag{25}$$

and the system

$$f(\lambda) = \lambda f(\mu) = \lambda\mu f(\theta/\lambda\mu n). \tag{26}$$

Of course, an upper bound for the usual codes $f$ employed in the above calculations is of key importance.

This suggests two ways of improving the upper bounds. One can either prove theorems stronger and more refined than Theorems 1–3 above or improve on the function $f$ for the usual error-correcting codes. Naturally, an improvement of bounds (21)–(23) obtained by employing already existing better functions $f$ cannot be counted as an achievement. Moreover, one of the simplest bounds, namely, the Plotkin bound, though it yields "bad" values, helps to visualize our results. Indeed, recall that for the Plotkin bound, $f$ has the form $1 - qd/n(q-1)$. Cancel out $\lambda$ and $\mu$ in (26) to obtain an explicit bound for the linear case. We get

$$(k/n)(1 + q/(q-1) + q^2/(q-1)^2) \leq 1 - (\theta/n)q^3/(q-1)^3. \tag{27}$$

For $q = 2$, this yields

$$k/n \leq (1 - 8\theta/n)/7, \tag{28}$$

which for $t = 0$ gives

$$k/n \leq 1/7. \tag{29}$$

In the same way, for the nonlinear case we easily find that

$$k/n \leq 1/3. \tag{30}$$

The Elias bound gives the values

$$k/n \leq .1263 \tag{31}$$

and

$$k/n \leq .3045, \tag{32}$$

respectively.

The Singleton bound for codes gives another explicit relation $\theta \leq n - 3(k-1)$, no less transparent than (27), whence

$$k/n \leq (1/3)(1 - \theta/n). \tag{33}$$

This expression is independent of $q$ and coincides with (27) if one assumes that $q \to \infty$ there. Recall the lower bound (14) to conclude that for $q \to \infty$, upper and lower bounds for linear $(2, 2, 2t)$-s.s. coincide.

All the upper bounds imply that for $q = 2$, the bound meets the $\theta/n$ axis at the point $(0, 1/8)$. This point was first obtained in [34].

Put $t = 0$ in (25) to obtain $f(\theta/\lambda n)$ and

$$f(\lambda) = \lambda \tag{34}$$

Thus, $\lambda$ is the point of an upper bound for error-correcting codes where the transmission rate equals the relative code distance. For $q = 2$, the Singleton, Plotkin, and Elias bounds give $1/2, 1/3$, and .3045, respectively. The last two values are the same as those in (30), (32).

Equation (34) was rediscovered in [19]. The authors of [19] did not improve the expression (25) itself. They rather employed the best known bound $f$ [29] to obtain

$$f(\lambda) = \lambda = .283477. \tag{35}$$

Unfortunately, they missed the next step, i.e., did not consider linear $(2, 2)$-s.s. Otherwise they would have arrived at the following system:

$$f(\lambda) = \lambda f(\mu) = \lambda\mu, \tag{36}$$

which is immediate from (26) for $t = 0$ and was already obtained in [47, 48]. The second equality in (26) implies $f(\mu) = \mu$. For $q = 2$, from (35) we get $\mu = .283477$ whence $f(\lambda) = \lambda 0.283477$. Again using [29], we obtain $\lambda = .381290$ and $f(\lambda) = .381290 \cdot .283477 = .108087$ in place of (31).

In [57], it is shown that the length of a code contained in a $(2, 2, 2t)$-c.s.s. is not more than half the code distance of this $(2, 2, 2t)$-s.s. It remains to apply any known upper bound for error-correcting codes to these two codes simultaneously. The most convenient is the Plotkin bound, which gives $k/n \le 1-2d/n$ and $k/n \le d/2-4\theta/n$, whence $k/n \le 1/5-16\theta/5n$. For $t = 0$, this gives $k/n \le 1/5$. The Elias bound and the bound in [29] for $t = 0$ give $k/n \le .1875$ and $k/n \le .171229$. For $k/n = 0$, all three bounds give $\theta \le n/16$. Recall that in the above numerical range, it is convenient to apply the bound in [29] in the form

$$k/n \le f(\lambda) = H_2((1 - \sqrt{1-(1-2\lambda)^2}/2).$$

Most of the argument just discussed was also published in [49, 50], and new values of the upper bounds were calculated in [58].

Note that in [19], the separating systems are studied in connection with the so-called "hash functions." The references therein include some related work by other authors [8, 9, 20, 21].

## 4. CONSTRUCTION OF $(2, 2, 2t)$-S.S.

We have already noted in the introduction that $(2, 2, 2t)$-s.s. were first constructed in [46] using Hadamard matrices. The idea of another construction (applying the MacDonald codes) has already been found in the work of Liu. However, he did not consider the problem of error correction and it is not advantageous to employ equidistant codes with large distance for the exceptional purpose of race-free coding, i.e., for $t = 0$. The next step was made in [34]. Namely, it was proved that in the codes constructed from Hadamard matrices of order $2^m, m \ge 2$, the quantity $\theta$ obeys the following relation:

$$\theta = \begin{cases} 2^{m-3}, & \text{if } m > 2 \\ 1 & \text{otherwise.} \end{cases} \tag{37}$$

For Hadamard matrices of order $4\ell = q + 1$, where $q = p^\alpha$, $p$ an odd prime and $\alpha$ a positive integer, it is shown in [36] that $[(q + 1)/8] \ge \theta \ge \max\{1, ](q - 2\sqrt{q} - 4)/9[\}$ and the lower bound is tight.

In [37] (and later, in [48]), it is shown that for binary linear $(2, 2, 2t)$-s.s. always

$$4D - 3d \ge 4\theta \ge 4d - 3D. \tag{38}$$

If the code is equidistant, i.e., $D = d$, Eq. (38) implies that $4\theta = d$ and (37) follows as a particular case. Equation (38) implies that the condition

$$4d - 3D > 0 \tag{39}$$

is sufficient for a linear code with minimum and maximum distances satisfying (39) to form a $(2, 2, 2t)$-s.s. with some $t > 0$. Equation (39) was used in [37,48, 49] to show that certain subcodes of the second order Reed–Muller codes form a $(2, 2, 2t)$-s.s. Namely, the codes of length $n = 2^m - 1$ with the number of information symbols

$$\begin{array}{llllll} k = 2m, & m \ge 7, & m \text{ odd}, & & & \\ k = 3m, & m \ge 9, & m \text{ odd}, & k = 3m/2, & m \ge 6, & m \text{ even} \\ k = 4m, & m \ge 11, & m \text{ odd}, & k = 5m/2, & m \ge 8, & m \text{ even} \\ k = 5m, & m \ge 17, & m \text{ odd}, & & & \\ k = 6m, & m \ge 23, & m \text{ odd}. & & & \end{array}$$

Condition (39) is related to the requirement that not only the minimum distance be sufficiently large but also that the maximum distance be sufficiently small. A.A. Nikanorov [36] constructed an example which shows that (39) is not necessary.

However, it is natural to look for $(2, 2, 2t)$-s.s. among the codes with a large code distance. In [48, 49], we have announced, and in [51] proved that almost all binary linear codes with code distance $d > .35n$ form $(2, 2, 2t)$-s.s.

Up to now in this section, we have discussed only binary $(2, 2, 2t)$-s.s. An important example of $(2, 2, 2t)$-s.s. with an alphabet of size $q > 2$ is provided by maximum length sequences ($n = q^m - 1$). As follows from [49, 52], in this case,

$$\begin{array}{l} \theta = 1, \text{ if } m = 2 \text{ and } q = 2, \\ \theta = 2, \text{ if } m = 2 \text{ and } q = 3, \\ \theta = q^{m-3}((q-1)^2 + (q-1)(q-2)^2), \text{ if } m \geq 3 \text{ and } q = 2 \text{ or } q = 3, \\ \theta = q - 1, \text{ if } m = 1 \text{ and } q > 3, \\ \theta = q^{m-2}(q-1)(q-3), \text{ if } m \geq 2 \text{ and } q > 3. \end{array} \tag{40}$$

Another efficient class of $(2, 2, 2t)$-s.s. with $q > 2$ is formed by maximum distance separable (MDS) codes [12, 18]. This was first shown in [53, 55]. Namely, for an MDS code,

$$\theta \geq n - 4(k-1) \tag{41}$$

and

$$\begin{array}{ll} \theta \leq n - 3(k-1) - 1 & \text{for } k > 1, \\ \theta = n & \text{for } k = 1. \end{array} \tag{42}$$

We see that for $k = 1$ and $k = 2$ the bounds (41) and (42) coincide.

In 1986, G.L. Katsman and S.N. Litsyn (unpublished) applied Mattson–Solomon polynomials and linearized polynomials [12] to Reed–Solomon (RS) codes (a class of MDS codes) to bring the bound (42) down to (41). Thus,

$$\theta = n - 4(k-1). \tag{43}$$

Comparing (40) with (41), (42), and (43), we conclude that MDS codes and especially Reed–Solomon codes are more efficient than the maximal length sequences.

Taking into account the equality $D = n$ valid for MDS codes, it is easy to derive from (41) and (42) the following relations:

$$\begin{array}{ll} 3d - 2D - 1 \geq \theta \geq 4d - 3D & \text{for } k > 1, \\ \theta = d = D & \text{for } k = 1. \end{array} \tag{44}$$

The second inequality in (44) has no factor 4 as compared to the second inequality in (38), and is valid only for MDS codes. Taking into account (43) we obtain for RS codes

$$\begin{array}{ll} \theta = 4d - 3D & \text{for } k > 1, \\ \theta = d = D & \text{for } k = 1. \end{array} \tag{45}$$

In the construction of $(2, 2, 2t)$-s.s. , the concept of concatenated codes plays an important role. Indeed, we can use a binary $(2, 2, 2t)$-s.s. with the parameters $n_1, k_1, \theta_1$ already constructed as the inner $(2, 2, 2t_1)$-s.s. and an MDS code (for instance, an RS code over $GF(2^m)$ with the parameters $n_2, k_2, \theta_2$) as the outer $(2, 2, 2t_2)$-s.s. to construct a super-$(2, 2, 2t)$-s.s. with the parameters $N = n_1n_2, K = k_1k_2, \theta \geq \theta_1\theta_2$.

The parameters of some inner and outer systems as well as the corresponding super-$(2, 2, 2t)$-s.s. are shown in Table 4.1. An asterisk marks s.s. obtained by puncturing $\theta_1 - i_1$ and $\theta_2 - i_2$ parity-check digits in s.s. of greater length located in the upper nearest line. The outer s.s. in lines 19 and 20 are formed by lengthening the s.s. of length 7 by 1 and 2 digits. In some lines $k_1 > m_1$. However, this does not violate the cascade code construction.

Table 4.1. Parameters of Cascade $(2, 2, 2t)$-s.s.

| N | Inner Binary s.s. | | | Outer s.s. over $GF(2^m)$ | | | | Super s.s. | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $n_1$ | $k_1$ | $\theta_1$ | $n_2$ | $k_2$ | $\theta_2$ | $m$ | $N$ | $K$ | $\theta$ |
| 1 | 7 | 3 | 1 | 7 | 2 | 3 | 3 | 49 | 6 | 3 |
| 2 | 7 | 3 | 1 | ∗6 | 2 | 2 | 3 | 42 | 6 | 2 |
| 3 | 7 | 3 | 1 | ∗5 | 2 | 1 | 3 | 35 | 6 | 1 |
| 4 | 15 | 4 | 2 | 15 | 3 | 7 | 4 | 225 | 12 | 14 |
| 5 | 15 | 4 | 2 | ∗14 | 3 | 6 | 4 | 210 | 12 | 12 |
| 6 | 15 | 4 | 2 | ∗13 | 3 | 5 | 4 | 195 | 12 | 10 |
| 7 | 15 | 4 | 2 | ∗12 | 3 | 4 | 4 | 180 | 12 | 8 |
| 8 | 15 | 4 | 2 | ∗11 | 3 | 3 | 4 | 165 | 12 | 6 |
| 9 | 15 | 4 | 2 | ∗10 | 3 | 2 | 4 | 150 | 12 | 4 |
| 10 | 15 | 4 | 2 | ∗9 | 3 | 1 | 4 | 135 | 12 | 2 |
| 11 | ∗14 | 4 | 1 | ∗9 | 3 | 1 | 4 | 126 | 12 | 1 |
| 12 | 15 | 4 | 2 | 15 | 4 | 3 | 4 | 225 | 16 | 6 |
| 13 | 15 | 4 | 2 | ∗14 | 4 | 2 | 4 | 210 | 16 | 4 |
| 14 | 15 | 4 | 2 | ∗13 | 4 | 1 | 4 | 195 | 16 | 2 |
| 15 | ∗14 | 4 | 1 | ∗13 | 4 | 1 | 4 | 182 | 16 | 1 |
| 16 | 15 | 4 | 2 | 7 | 2 | 3 | 3 | 105 | 8 | 6 |
| 17 | 15 | 4 | 2 | ∗6 | 2 | 2 | 3 | 90 | 8 | 4 |
| 18 | 15 | 4 | 2 | ∗5 | 2 | 1 | 3 | 75 | 8 | 2 |
| 19 | 7 | 3 | 1 | 8 | 2 | 4 | 3 | 56 | 6 | 4 |
| 20 | 7 | 3 | 1 | 9 | 3 | 1 | 3 | 63 | 9 | 1 |

Observe that if the inner $(2, 2, 2t)$-s.s. with distance $\delta$ is equidistant, it is easy to find the minimum and maximum distance for the super-s.s. Namely, we have $d = \delta(n - (k - 1))$ and $D = \delta n$. Substitute this into (38) and use the fact that in equidistant codes, $\theta = \delta/4$ to obtain $4\theta \geq \delta(n - 4(k - 1))$. On the other hand, $\theta_1 = \delta/4$ and $\theta_2 \geq n - 4(k - 1)$, whence $\theta \geq \theta_1\theta_2 \geq (\delta/4)(n - 4(k - 1))$. Thus, for this particular case, both ways of calculation give identical results.

The concept of generalized cascade codes [3], which had produced striking results for error-correcting codes, was not so efficient for $(2, 2, 2t)$-s.s.

Cascade codes were also of great use for the construction of $(2, 2, 2t)$-c.s.s. It is helpful that the outer code does not have to be a c.s.s. (it is sufficient if it is just an s.s.). Therefore, one can take linear codes, namely, MDS codes, as outer ones.

For $t = 0$, a table of parameters of $(2, 2, 2t)$-s.s. and methods of their construction are given in [10]. All of them yield small values of $k/n$. Some flaws of [10] were pointed out in [32]. In this paper, the authors also re-proved the main theorem of [41] and suggested an iterative construction method for $(2, 2, 2t)$-s.s. with improved parameters.

Thus, race-free and error-correcting coding of automata states form two faces of one and the same problem. The first part of this problem requires a much greater effort.

## 5. $(2, 1, 2t)$-S.S. BOUNDS AND CONSTRUCTION

There are models of asynchronous automata that impose weaker requirements on race-free coding as compared to those producing $(2, 1, 2t)$-s.s.

Such a model was first proposed in [68, 69]. Later, papers [13, 26, 27, 42, 43,] also suggested studying $(2, 1, 2t)$-s.s. Actually, the $(2, 1, 2t)$-s.s. can be obtained from $(2, 1, 2t)$-s.s. by putting $s_1 = s_2$ in (2). Then the coding problem has the following form. Construct a vector set of size $M$ such that any ordered triple of vectors

$$\begin{aligned} s_2 &= (a_1^{(2)}, a_2^{(2)}, \vdots, a_i^{(2)}, \vdots, a_n^{(2)}) \\ s_3 &= (a_1^{(3)}, a_2^{(3)}, \vdots, a_i^{(3)}, \vdots, a_n^{(3)}) \\ s_4 &= (a_1^{(4)}, a_2^{(4)}, \vdots, a_i^{(4)}, \vdots, a_n^{(4)}) \end{aligned} \tag{46}$$

contain at least $\theta^* \geq 2t + 1$ columns of the form

$$a_i^{(j)} \neq a_i^{(2)}, \quad j = 3, 4. \tag{47}$$

For $q = 2, t = 0$, Eqs. (46) and (47) were considered in [10] without involving any practical model of automata. In that paper, codes that satisfy these restrictions were constructed. Similarly to $(2, 2, 2t)$-s.s., linear $(2, 1, 2t)$-s.s. exist only for Model 2 of races.

In linear $(2, 1, 2t)$-s.s. , it is necessary to put $s_2 = 0$ and require that any pair of vectors

$$\begin{aligned} s_3 &= (a_1^{(3)}, a_2^{(3)}, \vdots, a_i^{(3)}, \vdots, a_n^{(3)}) \\ s_4 &= (a_1^{(4)}, a_2^{(4)}, \vdots, a_i^{(4)}, \vdots, a_n^{(4)}) \end{aligned} \tag{48}$$

contain at least $\theta^* \geq 2t + 2$ columns of the form

$$a_i^{(j)} \neq 0, \quad j = 3, 4. \tag{49}$$

It is clear that $(2, 1, 2t)$-s.s. form a particular case of $(2, 2, 2t)$-s.s. since any $(2, 1, 2t)$-s.s. is at the same time a $(2, 1, 2t)$-s.s. Therefore, all the results about $(2, 2, 2t)$-s.s. cited above can be easily reformulated for $(2, 1, 2t)$-s.s. (see [55]), which hardly requires any comment. In other words, all the results concerning $(2, 1, 2t)$-s.s. form simple corollaries of the above theory.

5.1. Existence bound.

$$k/n < (1/2)\{2 - (1 - \theta^*/n)\log_q(2q - 1) - 2(\theta^*/n)\log_q(q - 1) - H(\theta^*/n)\}. \tag{50}$$

The curve that bounds the domain (50) meets the axes $k/n$ and $\theta^*/n$ at the points $(\theta^*/n = 0, k/n = (1/2(2 - \log_q(2q - 1))$ and $(\theta^*/n = ((q - 1)/q^2, k/n = 0)$. This means that for $q = 2$, the axes intersect at the points $(0; (1/2)(2 - \log_2 3) = .207)$ and $(.25; 0)$. Asymptotically on $n$, the lower bounds for linear and nonlinear systems are identical despite the linear dependence of some pairs of vectors in (48).

Bound (50), similarly to (10), was found by a random coding argument. In 1987, S. N. Litsyn (unpublished) suggested the use of the bound for algebraic-geometric codes [65,67,], which for $q = p^{2m}, m$ prime, yields the following existence bound of $(2, 1, 2t)$-s.s.: $k/n \leq 1/2 - 1/(p^m - 1) - \theta^*/2n$.

For $m$ sufficiently large, this bound is better than (50). For instance, for $q = 1024$ and $t = 0$, it gives $k/n = .469697$, while (50) gives $k/n = .4500346$.

5.2. The following analog of Theorem 2 holds true. Any linear $(2, 1, 2t)$-s.s. contains a linear code of size $M$ whose length $d$ equals the minimum weight of the $(2, 1, 2t)$-s.s. and distance $\delta$ is at least $\theta^*$. This implies the following simultaneous inequalities:

$$k/n \leq f(\lambda); \quad k/n \leq \lambda f(\theta^*/\lambda n), \quad \lambda = d/n, \tag{51}$$

where, as before, $k/n \leq f(d/n)$ is any of the known upper bounds for the usual error-correcting codes. Notice that system (51) for linear $(2,2,2t)$-s.s. has the same form as for nonlinear $(2,2,2t)$-s.s.

The Singleton and Plotkin bounds yield, respectively,

$$\theta^* \leq n - 2(k-1), \tag{52}$$

$$(k/n)(1 + q/(q-1)) \leq 1 - (\theta^*/n)q^2/(q-1)^2. \tag{53}$$

At the point $k/n = 0, \theta^*/n = q^2/(q-1)^2$, bound (53) coincides with (50). The latter bound coincides with (51) and (52) asymptotically on $q$.

System (51) implies the equation $f(\lambda) = \lambda f(\theta^*/\lambda n)$. For $\theta^* = 1$, this equation turns into $f(\lambda) = \lambda$. The solutions to it for the Singleton, Plotkin, and Elias bounds, and for bound (52) are already known (from the above argument) to be equal to 1/2, 1/3, .3045, and .283477, respectively.

5.3. For $q = 2$, the following analog of inequalities (38) holds true:

$$2D - d \geq 2\theta^* \geq 2d - D. \tag{54}$$

These inequalities are also valid for nonlinear $(2,2,2t)$-s.s. The second of the inequalities yields the following sufficient condition that the code forms a $(2,1,2t)$-s.s.:

$$2d - D \geq 0 \tag{55}$$

(cf. (3)). For equidistant codes, (54) yields

$$\theta^* = d/2, \tag{56}$$

which implies that for codes of length $n = 4i-1 \geq 3, i = 1, 2, \vdots$ constructed from Hadamard matrices of order $n+1$, the quantity $\theta^*$ satisfies the equality $\theta^* = (n+1)/4$.

In a natural way, we can extend the list of subcodes of the second-order Reed–Muller codes that form $(2,1,2t)$-s.s.

For the maximal length sequences,

$$\theta^* = q^{m-2}(q-1)^2. \tag{57}$$

In other words, $\theta^* = d(q-1)/q$, which for $q = 2$ coincides with (56).

5.4. Analogously to $(2,2,2t)$-s.s. , almost all binary linear codes with $d \geq .2385n$ form $(2,1,2t)$-s.s.

5.5. For MDS codes,

$$\theta^* = n - 2(k-1), \tag{58}$$

and they form $(2,1,2t)$-s.s. with the maximal possible value of $\theta^*$ according to (52). Equation (58) implies $\theta^* = 2d - D$ since $D = n$.

5.6. As in the case of $(2,2,2t)$-s.s. above, MDS codes are of key importance for the construction of cascade $(2,1,2t)$-s.s. From (58) it follows that the choice of outer Reed–Solomon codes yields

$$k_2/n_2 = 1/2 \quad \text{for } \theta_2^* = 1. \tag{59}$$

Obviously, it is impossible to reach the bound in (50) with cascade $(2,1,2t)$-s.s. However, if one allows a small search in order to construct an inner $(2,1,2t)$-s.s. that meets this bound, it becomes possible to construct $(2, 1, 2t)$-s.s. with $K/N = .1$ for $\theta^* = 1$ and $q = 2$, which is better than the value .07 in [58]. For $\theta_1^* = 1$, this bound on $K/N$ follows from the fact that (50) gives $k_1/n_1 \geq .2$. It remains only to calculate $K/N = (k_1/n_1)(k_2/n_2)$.

Equations (57) and (58) produce many cascade $(2, 1, 2t)$-s.s. Some of them are listed in Table 5.1 [55, 59]. In this table, the letter "l" ("s") followed by figures, as, for example l1 (s1), marks lengthened (shortened) Reed–Solomon codes. The digit denotes the magnitude of lengthening (shortening). An asterisk marks inner $(2, 1, 2t)$-s.s. contained by puncturing $(2, 1, 2t)$-s.s. of greater length which appear in the upper nearest line. The table is based on the three inner $(2, 1, 2t)$-s.s. constructed from the maximal length sequences of length $n = 3, 7, 15$, and three Reed–Solomon codes of the same respective length over $Gf(2^2), GF(2^3)$, and $GF(2^4)$. Moreover, the inner $(2, 1, 2t)$-s.s. with $n_1 = 9, k_1 = 4$, and $\theta_1^* = 1$ is the casecade $(2, 1, 2t)$-s.s. constructed in the first line of Table 5.1. The tables of $(2, 1, 2t)$-s.s. constructed by shortening, puncturing, and from Eq. (54) (see Table 5.2) are given in [42, 43]. A comparison with these tables shows that up to length 135 the s.s. in Table 5.1 fill the length range more densely, the length range itself is broader, and that many s.s. from Table 5.1 are not listed in [42, 43]. It is an easy matter to expand Table 5.1 by longer s.s. and to make it equally "dense." Moreover, one can expand Table 5.1 building on its own s.s. For example, combine the inner $(2, 1, 2t)$-s.s. in line 18 and the outer $(2, 1, 2t)$-s.s. in line 10 to construct a $(2, 1, 2t)$-s.s. with $N = 63, K = 16, \theta^* = 1$. The code distance of this super-$(2, 1, 2t)$-s.s. is easily computed to be $d = 16$. In [43], Pradhan also gives a table of $(2, 1, 2t)$-s.s. obtained as the direct product of a $(2, 1, 2t)$-s.s. onto itself and onto the codes in Table 5.2.

The advantages of s.s. in Table 5.1 follow from a more flexible construction of cascade $(2, 1, 2t)$-s.s. When the length becomes large, the capacities of this construction show themselves even better. For instance, take three cascade $(2, 1, 2t)$-s.s. constructed from one and the same inner $(2, 1, 2t)$-s.s. with the parameters $n_1 = 15, k_1 = 6, \theta_1^* = 1$. This system is a super-$(2, 1, 2t)$-s.s. shown in line 3 of Table 5.1. Take three separating systems with the parameters $n_2 = 19, k_2 = 7, \theta_2^* = 7; n_2 = 31, k_2 = 15, \theta_2^* = 3; n_2 = 53, k_2 = 19, \theta_2^* = 17$ as the outer ones. The first and the third systems are obtained from RS codes with the parameters $n_2 = 31, k_2 = 19; n_2 = 63, k_2 = 29$ by a shortening by 12 and 10, respectively, while the second system is itself an RS code. The super-$(2, 1, 2t)$-s.s. will have the parameters $N = 258, K = 42, \theta^* = 7; N = 465, K = 90, \theta^* = 3; N = 795, K = 114, \theta^* = 17$, respectively. For comparison, we gave the following three $(2, 1, 2t)$-s.s. from [43]: $N = 378, K = 42, \theta^* = 7; N = 762, K = 88, \theta^* = 3; N = 1530, K = 114, \theta^* = 17$. These systems have much greater length, while the remaining parameters are identical.

Almost all of the results concerning the $(2, 1, 2t)$-s.s. as a particular case of $(2, 2, 2t)$-s.s. obtained in [53–55] were later collected in [59].

In [44], the $(2, 2, 2t)$-s.s. and $(2, 1, 2t)$-s.s. are constructed from Berger's codes [2].

Finally, it remains to consider the $(2, 1, 2t)$-c.s.s., i.e., a set of vectors over $GF(2)$ of size $M$ such that in any ordered triple (46), there exist at least $\theta^* \geq 2t + 1$ columns of both forms $(011)^T$ and $(100)^T$ at the same time.

The existence bound for $q = 2$ has the following form [57]:

$$R < (1/2)(3 - H(\theta^*/n) - (1 - \theta^*/n)\log_2 7),$$

whence $R = .0963$ for $\theta^*/n = 0$ and $\theta^*/n = 1/8$ for $R = 0$.

We were not able to prove that $(2, 1, 2t)$-c.s.s. have the same property as $(2, 2, 2t)$-s.s. which allows one to derive an upper bound, namely, that a code of length $n$ contains a code of length $\delta$ or $n - \delta$ with distance expressed via $\theta^*$. The only thing that can be proved for $(2, 1, 2t)$-c.s.s. [57] is that $d \geq 2\theta^*$ and $n - d \geq 2\theta^*$, which implies $n \geq 4\theta^*$. In [4], it is shown that linear $(2, 1, 2t)$-c.s.s. and $(2, 2, 2t)$-c.s.s. do not exist. Constructions of c.s.s. were suggested in [28, 44]. However, in those papers, the length $n$ is of order $(\log_2 M)^\alpha$, where $\alpha \geq 2$. This means that the redundancy of these

Table 5.1. Parameters of Cascade $(2, 2, 2t)$-s.s.

| N | Inner Binary s.s. | | | Outer s.s. over $GF(2^m)$ | | | | Super s.s. | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $n_1$ | $k_1$ | $\theta_1$ | $n_2$ | $k_2$ | $\theta_2$ | $m$ | $N$ | $K$ | $\theta$ |
| 1 | 3 | 2 | 1 | 3 | 2 | 1 | 2 | 9 | 4 | 1 |
| 2 | 3 | 2 | 1 | 4(l1) | 2 | 2 | 2 | 12 | 4 | 2 |
| 3 | 3 | 2 | 1 | 5(l2) | 3 | 1 | 2 | 15 | 6 | 1 |
| 4 | 7 | 3 | 2 | 3 | 2 | 1 | 2 | 21 | 6 | 2 |
| 5 | 7 | 3 | 2 | 4(l1) | 2 | 2 | 2 | 28 | 6 | 4 |
| 6 | ∗6 | 3 | 1 | 5(l2) | 3 | 1 | 2 | 30 | 9 | 1 |
| 7 | ∗6 | 3 | 1 | 5(l2) | 3 | 2 | 4 | 30 | 6 | 3 |
| 8 | 7 | 3 | 2 | 5(l2) | 2 | 3 | 2 | 35 | 6 | 6 |
| 9 | 7 | 3 | 2 | 5(l2) | 3 | 1 | 2 | 35 | 9 | 2 |
| 10 | ∗6 | 3 | 1 | 7 | 4 | 1 | 3 | 42 | 12 | 1 |
| 11 | 7 | 3 | 2 | 6(s1) | 3 | 2 | 3 | 42 | 9 | 4 |
| 12 | 7 | 3 | 2 | 6(s1) | 2 | 4 | 3 | 42 | 6 | 8 |
| 13 | ∗6 | 3 | 1 | 8(l1) | 4 | 2 | 3 | 48 | 12 | 2 |
| 14 | 7 | 3 | 2 | 7 | 3 | 3 | 3 | 49 | 9 | 6 |
| 15 | 7 | 3 | 2 | 7 | 2 | 5 | 3 | 49 | 6 | 10 |
| 16 | ∗6 | 3 | 1 | 9(l2) | 5 | 1 | 3 | 54 | 15 | 1 |
| 17 | ∗6 | 3 | 1 | 9(l2) | 4 | 3 | 3 | 54 | 12 | 3 |
| 18 | 9 | 4 | 1 | 8(l1) | 4 | 2 | 3 | 72 | 16 | 2 |
| 19 | 9 | 4 | 1 | 9(l2) | 4 | 3 | 3 | 81 | 16 | 3 |
| 20 | 9 | 4 | 1 | 9(l2) | 5 | 1 | 3 | 81 | 20 | 1 |
| 21 | 9 | 4 | 1 | 9)l2 | 3 | 5 | 3 | 81 | 12 | 5 |
| 22 | 9 | 4 | 1 | 10(s5) | 4 | 4 | 4 | 91 | 16 | 4 |
| 23 | 15 | 4 | 4 | 8(s1) | 3 | 2 | 3 | 90 | 12 | 8 |
| 24 | ∗14 | 4 | 3 | 7 | 3 | 3 | 3 | 98 | 12 | 9 |
| 25 | ∗13 | 4 | 2 | 7 | 3 | 3 | 3 | 91 | 12 | 6 |
| 26 | 9 | 4 | 1 | 11(s1) | 4 | 5 | 4 | 99 | 16 | 5 |
| 27 | 15 | 4 | 4 | 7 | 3 | 3 | 3 | 105 | 12 | 12 |
| 28 | 9 | 4 | 1 | 12(s3) | 4 | 6 | 4 | 108 | 16 | 6 |
| 29 | 9 | 4 | 1 | 13(s2) | 7 | 1 | 4 | 117 | 28 | 1 |
| 30 | 9 | 4 | 1 | 13(s2) | 6 | 3 | 4 | 117 | 24 | 3 |
| 31 | 9 | 4 | 1 | 13(s2) | 5 | 5 | 4 | 117 | 20 | 5 |
| 32 | 9 | 4 | 1 | 14(s1) | 7 | 2 | 4 | 126 | 28 | 2 |
| 33 | 9 | 4 | 1 | 14(s1) | 6 | 4 | 4 | 126 | 24 | 4 |
| 34 | ∗14 | 4 | 3 | 9(l2) | 3 | 5 | 3 | 126 | 12 | 15 |
| 35 | 9 | 4 | 1 | 15 | 8 | 1 | 4 | 135 | 32 | 1 |
| 36 | 9 | 4 | 1 | 15 | 7 | 3 | 4 | 135 | 28 | 3 |
| 37 | 9 | 4 | 1 | 15 | 6 | 5 | 4 | 135 | 24 | 5 |
| 38 | 9 | 4 | 1 | 15 | 5 | 7 | 4 | 135 | 20 | 7 |
| 39 | 15 | 4 | 4 | 14(s1) | 7 | 2 | 4 | 210 | 28 | 8 |
| 40 | 15 | 4 | 4 | 15 | 8 | 1 | 4 | 225 | 32 | 4 |
| 41 | 15 | 4 | 4 | 16(l1) | 8 | 2 | 4 | 240 | 32 | 8 |
| 42 | 15 | 4 | 4 | 16(l1) | 7 | 4 | 4 | 240 | 28 | 16 |
| 43 | 15 | 4 | 4 | 17(l2) | 8 | 3 | 4 | 255 | 32 | 12 |

Table 5.2. Parameters of $(2, 1, 2t)$-s.s.

| $N$ | $n$ | $k$ | $\theta^*$ | $N$ | $n$ | $k$ | $\theta^*$ |
|---|---|---|---|---|---|---|---|
| 1 | 12 | 4 | $\geq 1$ | 8 | 14 | 4 | 3 |
| 2 | 30 | 10 | $\geq 1$ | 9 | 30 | 5 | 7 |
| 3 | 60 | 15 | $\geq 1$ | 10 | 62 | 15 | 3 |
| 4 | 126 | 28 | $\geq 1$ | 11 | 62 | 9 | 9 |
| 5 | 254 | 44 | $\geq 1$ | 12 | 126 | 21 | 7 |
| 6 | 510 | 75 | $\geq 1$ | 13 | 126 | 14 | 17 |
| 7 | 1020 | 120 | $\geq 1$ | 14 | 254 | 36 | 9 |
| | | | | 15 | 254 | 29 | 15 |

codes tends to 1. It should be noted that in [28] it was conjectured that the ratio of the minimum lengths of the corresponding c.s.s. and s.s. tends to 1 as the code size $M$ grows. However, this sounds plausible only for large values of $\theta/n$ and seems dubious for small ones. In the latter case, it seems reasonable to rely upon a construction method for c.s.s. from s.s. from this paper which suggests that one first construct an s.s. and then add inversions to obtain a c.s.s., thus doubling the length.

Concerning cascade $(2, 1, 2t)$-c.s.s. one should repeat the argument in the end of Section 4.

The problems related to separating systems have gradually lost the connection to their origin, namely, the coding of automata states. The latest papers devoted to s.s. and related to the "automata" trend ([17, 57, 59]) appeared or obtained a solution at the beginning of the eighties. Gradually the research in this area received an entirely new motivation. It turned out that linear $(2, 1, 2t)$-s.s. are useful for the construction of the so-called 3-covering systems. A system of $M$ binary vectors of length $n$ with the property that any $t$ coordinates contain all possible $2^t$ binary strings is called a binary $t$-covering system. (The construction problem of $t$ covering systems originates, in particular, in technical diagnose.)

Owing to a new problem source, the area began to involve another circle of researchers. In large part, they are isolated from the researchers of earlier times and assume that they deal with *tabula rasa*. They inevitably rediscover the known results. A new interest in s.s. started in the beginning of eighties and continues until now. A concise survey on linear $(2, 1, 2t)$-s.s., far from being complete, appears in [62]. This encompasses mostly the works of the last decade. It is true that the term "linear $(2, 1, 2t)$-s.s." is not used there. However, the term "intersecting code" appearing in the title of [65] is entirely equivalent to it. Apart from this, [62] does not require the condition $t > 0$. On intersecting codes, see [6, 16, 24, 31, 45].

Let us list the results in [62] explicitly pointing out the source: $d \geq k$ [24]. If $D < 2d$, then a linear code forms a $(2, 1, 2t)$-s.s. [6] (compare with (3) and (55)). If $[AI]$ is a generating matrix of a linear $(2, 1, 2t)$-s.s. with the parameters $n, k$, and $d$, then the matrix

$$\begin{Vmatrix} A & I & I & 0 \\ 00\ldots0 & 11\ldots1 & 00\ldots0 & 1 \end{Vmatrix}$$

generates a linear $(2, 1, 2t)$-s.s. with $n' = n + k + 1, k' = k + 1, d' = k + 1$ [38]. A direct product of two linear $(2, 1, 2t)$-s.s. with the parameters $n_1, k_1, d_1$ and $n_2, k_2, d_2$ is again a linear $(2, 1, 2t)$-s.s. with $n = n_1 n_2, k = k_1 k_2, d = d_1 d_2$ [38]. Linear systems from direct products are also constructed in [59].

If in an MDS code, $n = 2k - 1$, it forms a linear $(2, 1, 2t)$-s.s. (In fact, one should write $n \geq 2k - 1$; cf (58). In a binary $(2, 1, 2t)$-s.s., for any two nonzero vectors, there exists a coordinate in which one vector is 0 and the other is 1 [5]. This is immediate from the definition of $(2, 1, 2t)$-s.s. In [62], Sloane states the property of vector triples in linear $(2, 1, 2t)$-s.s., which is used for the construction of 3-covering systems. It seems appropriate to ask which $(i, j)$-s.s. can be used for the construction of 3-covering systems for different values of $t$. And, if the linearity of $(i, j)$-s.s. is essential, how does one handle the fact that under any of the two inequalities $i > 2, j > 2$ binary linear $(i, j)$-s.s. do not exist [56]?

In complete accordance with Sections 5.6 and 5.7 (see [53, 55]), Sloane [62] asserts that the best example of $(2, 1, 2t)$-s.s. is provided by the MDS codes. In fact, MDS codes, as stressed in (52) and (58), form $(2, 1, 2t)$-s.s. with the maximum possible value of $\theta^*$. Recall that they give at the same time the best example of $(2, 2, 2t)$-s.s., as follows from Section 4 (see its part following Eq. (41)).

Asymptotic bounds on the parameters of $(2, 1, 2t)$-s.s. are attributed to [6, 16, 45] (and also to an unpublished work by A. Blokhuis and Metsch). They form a complete replica of the bounds in

Sections 5.1 and 5.2 obtained earlier (see [55, 59] An important result was obtained in [61]. Namely, it is shown that the existence bound is attained on Goppa codes [14].

New separating system are investigated in [60, 63]. Interesting comparison of the bounds received in [7, 19, 39, 47], is submitted in [15]. The connection between (2, 2)- s.s. and superimposed codes is marked in [18].

The problems in separating systems had such a variety of sources that they started to form a self-contained mathematical area. Separating systems now form an independent object of mathematical research.

## 6. ON THE NEW POSSIBLE APPLICATION OF SEPARATING SYSTEMS

If we take an optional panel of DNA molecules, it becomes clear that it is a separating system. If it is not, it meant, that all the DNA are equal. If it is so, we can apply all the aspects which are applied to separating systems in general to this panel. After such analyze is completed we can make conclusions regarding quantitative relation, typical to DNA panels, probably in compliance with the borders of the parameters, presented in previous parts. Primary analyze of this kind is presented in this part.

**Glossary:**

**Nucleotide:** considered simple element. There is four of them: A (adenine), T (thymine), G (guanine), C (cytosine).

**Codon:** a bloc of three nucleotides. It is evident, that there are $4^3 = 64$ different codons. DNA consider as a consequence of codons.

**Amino acid:** as well as codon - a bloc of three nucleotides. There are 21 amino acids. It means that one or more codons correspond to one amino acid. See below a table 6.1 of amino acids and codons correspondence.

Table 6.1

| | Name | Symbol | Codons |
|---|---|---|---|
| 1 | Alanine | Ala | GCT,GCC,GCA,GCG |
| 2 | Arginine | Arg | CGT,CGC,CGA,CGG,AGA,AGG |
| 3 | Aspartic acid | Asp | GAT,GAC |
| 4 | Asparagine | Asn | AAT,AAC |
| 5 | Cysteine | Cys | TGT,TGC |
| 6 | Glutamine | Gln | CAA,CAG |
| 7 | Glutamic acid | Glu | GAA,GAG |
| 8 | Glycine | Gly | GGT,GGC,GGA,GGG |
| 9 | Histidine | Hig | CAT,CAC |
| 10 | Isoleucine | Ile | ATT,ATC,ATA |
| 11 | Leucine | Leu | TTA,TTG,CTT,CTC,CTA,CTG |
| 12 | Lysine | Lys | AAA,AAG |
| 13 | Methionine | Met | ATG |
| 14 | Phenylalanine | Phe | TTT,TTC |
| 15 | Proline | Pro | CCT,CCC,CCA,CCG |
| 16 | Serine | Ser | TCT,TCC,TCA,TCG,AGT,AGC |
| 17 | Threonine | Thr | ACT,ACC,ACA,ACG |
| 18 | Tryptophan | Trp | TGG |
| 19 | Tyrosine | Tyr | TAT,TAC |
| 20 | Valine | Val | GTT,GTC,GTA,GTG |
| 21 | Stop codon | STOP | TAA,TAG,TGA |

**DNA of the organisms researched**

In our work we researched DNA of the following organisms:

1. **Candida Glabrata:** Thrush fungus (more frequent agent, provoking scab to the humans).
2. **Candida Tropicalis:** The yeast Candida tropicalis is the second most pathogenic Candida species after Candida albicans and is more often associated with deep fungal infections than normal mucosa. Candida tropicalis is an asexual diploid organism. Similar to many other Candida species, a CUG codon in Candida tropicalis corresponds to a serine residue instead of the universal leucine. Candida tropicalis is used in industry for the preparation of polyester, polyamide, and perfume, and the formation of xylitol, a sugar alcohol that can replace sucrose. It is also an important organism for studying peroxisome biogenesis and peroxisomal protein expression. The exact genome size and chromosome number of Candida tropicalis are unknown, but it has been estimated to have a haploid genome size of 15 Mb, organized in 5 or 6 chromosomes.
3. **Debaryomyces Hansenii:** Trush fungus resistible to sea - salt. One can find it in cheese and fish. Non pathogenic, but very close to Candida Glabrata which is pathogenic.
4. **Encephalitozoon Cuniculi:** Has the shortest genome from the known eukaryotes. One celled fungus. Infects mammalian. May infected nervous system, respirator system and intestinal.
5. **Eremothecium Gossypii:** This fungus has the smallest genome of all independent eukaryotes researched. Cotton pathogen, spread by sucking insects.
6. **Gibberella Zeae:** This fungus, which damages grain varieties damaged USA agriculture. It also emits vomitoxin, which is a health hazard.
7. **Kluyveromyces Lactis:** Kluyveromyces lactis is a petite-negative hemiascomycete yeast. Compared to Saccharomyces cerevisiae, it can use a wider variety of carbon sources, and many of its strains were originally isolated from milk-derived products in which the major carbon source is lactose. Kluyveromyces lactis is commonly used in genetic research and for industrial applications, such as the production of beta-galactosidase and the heterologous proteins calf prochymosin, human serum albumin, and human interleukin-1-gamma. The Kluyveromyces lactis genome is approximately 10.6 Mb, organized in 6 chromosomes.
7. **Magnaporthe Grisea:** Magnaporthe grisea (anamorph Pyricularia grisea), a haploid filamentous Ascomycete, is the causal agent of rice blast disease. This worldwide rice disease is a major threat, destroying enough rice annually to feed more than 60 million people. Crop losses associated with this disease have been magnified in recent times with the intensification of rice production. Although resistant strains of rice have been developed, Magnaporthe grisea can rapidly evolve to overcome host resistance. Aside from rice, certain strains of Magnaporthe grisea are able to attack barley, wheat, pearl millet, and turfgrass. Magnaporthe grisea is an ideal model organism for studying plant pathogenic fungi and host-parasite interactions for several reasons: it has a relatively small genome, making it amenable to whole genome analysis; extensive genetic mapping data is available; it is closely related to the widely studied non-pathogen Neurospora crassa, enabling comparative genomic studies; a draft sequence of the host (rice) genome has been completed; it can be cultured on defined media and has a well-established transformation system, facilitating biochemical and molecular analyses; and the early stages of its infection process can be experimentally studied. A full understanding of the molecular bases of fungal phytopathogenicity and host-parasite interactions will be instrumental for the development of novel environmentally sound strategies to protect world food supplies. The Magnaporthe grisea genome is approximately 40 Mb, organized in 7 chromosomes.
8. **First human chromosome**
9. **Second human chromosome**

The data was taken from the data base of National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/)

**General statistic results** In this part we are going to check the hypothesis about nucleotide frequency equality (A, T, G, C), the nucleotides, which compose DNA sequences.
It means, that P(A)=P(T)=P(G)=P(C)=0.25.

Table 6.2. Nucleotide statistics for different DNA sequences:

| \| | DNA | DNA length | Nucleot. A | Nucleot. T | Nucleot. G | Nucleot. C |
|---|---|---|---|---|---|---|
| 1 | Cand. Glabr. | 12 937 596 | 0.306954 | 0.306861 | 0.193299 | 0.192886 |
| 2 | Cand. Tropic. | 14 690 277 | 0.333866 | 0.333746 | 0.166465 | 0.165923 |
| 3 | Debaryo. Hansen. | 12 335 868 | 0.318344 | 0.318447 | 0.181938 | 0.181241 |
| 4 | Encephal. Cunic. | 2 549 892 | 0.263273 | 0.262551 | 0.239261 | 0.234915 |
| 5 | Eremoth. Gossyp. | 8 923 452 | 0.241331 | 0.240738 | 0.259679 | 0.258253 |
| 6 | Kluyver. Lactis | 11 607 900 | 0.307079 | 0.30595 | 0.193746 | 0.193225 |
| 7 | Magna. Grisea | 38 851 788 | 0.241684 | 0.241611 | 0.258093 | 0.258612 |
| 8 | Homo Sapiens | 624 833 964 | 0.294715 | 0.294717 | 0.205044 | 0.205523 |
|  | AVERAGE |  | 0.2910676 | 0.291023295 | 0.2087749 | 0.209134 |

Taking into account the dates we've got we can make a guess about the frequencies of any kind of codons and amino acids appearance (See Table 6.3 and 6.4)

There are measured and calculated frequencies of amino acids in the table 3 of appendix. Measures were made using average frequencies of nucleotides.

Separating parameter for the pair sequences is similar to Hamming distance between DNA sequences

Theoretically calculated average meaning of the possibility of coincidence of two codes in different sequences equals 0.05667.

In the second part of our experiment we research the distance between different sequences. Distance here is a simple code distance. Three nucleotides sequence, amino acid is considered as a word.

So, we find out the difference between real and supposed distances, relying on the frequencies table, received at the first stage of the experiment.

Table 6.3

| Average frequency for codons of all dna sequences |  |  |
|---|---|---|
| AAA | 0.024659336 | 0.036949586 |
| AAC | 0.017717921 | 0.014896717 |
| AAG | 0.017687476 | 0.019815885 |
| AAT | 0.024655587 | 0.024609515 |
| ACA | 0.017717921 | 0.019710818 |
| ACC | 0.012730461 | 0.011994931 |
| ACG | 0.012708586 | 0.003611596 |
| ACT | 0.017715227 | 0.015739381 |
| AGA | 0.017687476 | 0.02158792 |
| AGC | 0.012708586 | 0.013972354 |
| AGG | 0.012686749 | 0.017141746 |
| AGT | 0.017684787 | 0.015700218 |
| ATA | 0.024655587 | 0.020320493 |
| ATC | 0.017715227 | 0.01390673 |
| ATG | 0.017684787 | 0.018168597 |
| ATT | 0.024651838 | 0.024629412 |
| CAA | 0.017717921 | 0.019467203 |

| Average frequency for codons of all dna sequences | | |
|---|---|---|
| CAC | 0.012730461 | 0.014723792 |
| CAG | 0.012708586 | 0.01959126 |
| CAT | 0.017715227 | 0.018191805 |
| CCA | 0.012730461 | 0.018367797 |
| CCC | 0.009146933 | 0.013030963 |
| CCG | 0.009131216 | 0.003851772 |
| CCT | 0.012728525 | 0.017182899 |
| CGA | 0.012708586 | 0.00360973 |
| CGC | 0.009131216 | 0.003491927 |
| CGG | 0.009115526 | 0.003850335 |
| CGT | 0.012706654 | 0.0035938 |
| CTA | 0.017715227 | 0.01262618 |
| CTC | 0.012728525 | 0.016438611 |
| CTG | 0.012706654 | 0.019617548 |
| CTT | 0.017712533 | 0.019816397 |
| GAA | 0.017687476 | 0.019774741 |
| GAC | 0.012708586 | 0.009909526 |
| GAG | 0.012686749 | 0.016438157 |
| GAT | 0.017684787 | 0.013879344 |
| GCA | 0.012708586 | 0.014451277 |
| GCC | 0.009131216 | 0.012071234 |
| GCG | 0.009115526 | 0.003480212 |
| GCT | 0.012706654 | 0.013964003 |
| GGA | 0.012686749 | 0.015305578 |
| GGC | 0.009115526 | 0.012049199 |
| GGG | 0.009099863 | 0.012937536 |
| GGT | 0.01268482 | 0.011937955 |
| GTA | 0.017684787 | 0.011485989 |
| GTC | 0.012706654 | 0.009894326 |
| GTG | 0.01268482 | 0.014691543 |
| GTT | 0.017682098 | 0.014839374 |
| TAA | 0.024655587 | 0.020048123 |
| TAC | 0.017715227 | 0.01153813 |
| TAG | 0.017684787 | 0.012593589 |
| TAT | 0.024651838 | 0.020343693 |
| TCA | 0.017715227 | 0.019442596 |
| TCC | 0.012728525 | 0.015346227 |
| TCG | 0.012706654 | 0.003625599 |
| TCT | 0.017712533 | 0.021606216 |
| TGA | 0.017684787 | 0.019415384 |
| TGC | 0.012706654 | 0.014443417 |
| TGG | 0.01268482 | 0.018320733 |
| TGT | 0.017682098 | 0.019681938 |
| TTA | 0.024651838 | 0.020076652 |
| TTC | 0.017712533 | 0.019796554 |
| TTG | 0.017682098 | 0.019443467 |
| TTT | 0.024648089 | 0.03695977 |

Table 6.4

| Average frequency for amino acids of all dna sequences | | |
|---|---|---|
| Ala | 0.043661982 | 0.043849795 |
| Arg | 0.074036208 | 0.05335044 |
| Asn | 0.042373508 | 0.039622307 |
| Asp | 0.030393373 | 0.023871536 |
| Cys | 0.030388752 | 0.03408558 |
| Gln | 0.030426507 | 0.039032859 |
| Glu | 0.030374226 | 0.036183784 |
| Gly | 0.043586958 | 0.051993098 |
| Hig | 0.030445688 | 0.032866478 |
| Ile | 0.067022652 | 0.05908309 |
| Leu | 0.103196875 | 0.108005897 |
| Lys | 0.042346813 | 0.056721514 |
| Met | 0.017684787 | 0.018174642 |
| Phe | 0.042360623 | 0.056727477 |
| Pro | 0.043737136 | 0.052193965 |
| Ser | 0.091256313 | 0.089710005 |
| STOP | 0.060025161 | 0.052125514 |
| Thr | 0.060872195 | 0.051121175 |
| Trp | 0.01268482 | 0.018269699 |
| Tyr | 0.042367065 | 0.032030794 |
| Val | 0.060758359 | 0.050980351 |

## 7. THE CONCLUSION

In given article the first step of possible application of the theory of separating systems in researches on genetics is designated only. In opinion of experts the further researches can be rather fruitful.

## REFERENCES[1]

1. Armstrong D. B., A general method of applying error correction to synchronous digital system. *Bell System Tech. J.,* 1961, 40, pp. 567–593.
2. Berger J. M. A note on error detection codes for asymmetric channels, *Inf. Contr.*, 1961, 4, pp. 68–73.
3. Blokh E. L. and Zyablov V. V. *Generalized Cascade Codes*, M.: Svyaz', 1976.
4. Bose B. and Rao T.R.N., Separating and completely separating systems and linear codes. *IEEE Trans. Comput.*, 1980, 29, pp. 665–668.
5. Busschbach P., Constructive Methods to Solve the Problems of $s$-Surjectivity, Conflict Resolution, Coding in Defective Memories, 1984, Rep. 84D005, ENST, Paris (1984).
6. Cohen G. D. and Lempel A.. Linear intersecting codes. *Discrete Math.*, 1985, 56, pp. 35–43.
7. Cohen G. D. and Zemor G. Intersecting Codes and Independent Families, *IEEE Trans. Inform. Theory*, 1994, 40, pp. 1872–1881.
8. Erdös P. F. and Füredi Z. Families of finite sets in which no set is covered by the union of two others. *J. Comb. Theory*, *Ser. A*, 1982, 33, pp. 158–166.

[1] All articles published in *Probl. Peredachi Inf.* are translated in English.

9. Fredman M. and Komlos J. On the size of separating systems and hash functions. *SIAM J. Alg. Discr. Methods*, 1984, 5, pp. 61–68.

10. Friedman A. D., Graham R. L. and Ulman J. D. Universal single transition time asynchronous state assignments. *IEEE Trans. Comput.*, 1969, 18, pp. 541–547.

11. Gavrilov M. A. Structural redundancy and reliability of relay devices. in: *IFAC First Int. Congress* (in Russian). M.: 1960, Izd. Akad. Nauk SSSR.

12. Gavrilov M. A., Ostianu V. M. and Potekhin A. I. Reliability of discrete devices. in: *Probability Theory. Mathematical Statistics. Technical Cybernetics* (in Russian), M.: 1970, VINITI.

13. Glebskii Yu. V., On the stability of asynchronous automata. *Avtomat. Telemekh.*, 1976, 12, pp. 114–119.

14. Goppa V. D. Rational representation and $(L, g)$-codes. *Probl. Peredachi Inf.*, 1981, 3, pp. 41–49.

15. Kabatiansky G. A., On Pair-Separatimg Codes, *Probl. Peredachi Inf.*, 2001, 4, pp. 60–62.

16. Katona G. O. H. and Srivastava, J. Minimal 2-coverings of a finite affine space based on $GF(2)$. *J. Stat. Plan. Inference*, 8, 1983, pp. 375–378.

17. Katsman G. L. and Litsyn S. N., Double fault-tolerant separating systems, *Automat. Telemekh.*, 1986, 11, pp. 114–117.

18. Kim H. K. and Lebedev V. S., On the Optimality of Trivial $(w, r)$ Cover-Free Codes,*Probl. Peredachi Inf.*, 2004, 40, 3, pp. 13–20.

19. Körner J. and Simonyi G., Separating partition systems and locally different sequences, *SIAM J. Discr. Math*, 1988, 1, pp. 355–359.

20. Körner J. and Marton K., New bounds for perfect hashing via information theory, *Eur. J. Comb*, 1988, 9, pp. 523–530.

21. Körner J., Fredman–Komlos bounds and information theory, *SIAM J. Alg. Discr. Methods*, 1986, 7, pp. 560–570.

22. Kuhl J. G. and Reddy S. M., A multicode single transition time state assignment for asynchronous sequential machines, *IEEE Trans. Comput.*, 1978, 2, pp. 927–934.

23. Lazarev V. G. and Pijl' E. I., *Synthesis of Control Automata* [in Russian], M.: Energiya, 1978.

24. Lempel A. and Winograd S., A new approach to error-correcting codes, *IEEE Trans. Inform. Theory*, 1977, 23, pp. 503–508.

25. Liu C. N., A state variable assignment method for asynchronous sequential switching circuits, *J. ACM*, 1963, 10, pp. 209–216.

26. Mago G., Asynchronous sequential circuits with $(2, 1)$-type state assignments, in: *IEEE Conf. Rec. 11th Annual Symp. Switch. and Automata Theory* New York 1970, pp. 109–113.

27. Mago G. Realization methods for asynchronous sequential circuits, *IEEE Trans. Comput.*, 1971, 20, pp. 290–297.

28. Mago G., Monotone functions in sequential circuits, *IEEE Trans. Comput.*,1973, 22, pp. 928–933.

29. McEliece R. J., Rodemich E. R., Rumsey H. C. and Welch L. R., New upper bounds on the rate of a code via Delsarte–MacWilliams inequalities, *IEEE Trans. Inf. Theory*, 1977, 23, pp. 157–166.

30. MacWilliams F. J. and Sloane N. J. A., *Error-Correcting Codes*, North Holland, 1977.

31. Miklòs D. Linear binary codes with intersection properties, *Discrete App. Math.*, 1984, 9, pp. 187–196.

32. Nanya T. and Tohma Y., On universal single transition time asynchronous state assignments, *IEEE Trans. Comput.*, 1978, 27, pp. 781–782.

33. Nanya T. and Tohma Y., Universal multicode STT assignments for asynchronous sequential machines. *IEEE Trans. Comput.*, 1979, 28, pp. 811–818.

34. Nemsadze N. K., State encoding of finite automata. *Probl. Peredachi Inf.*, 1969, 5, 1, pp. 79–86.

35. Nemsadze N. K. An application of equidistant codes to state encoding of automata, *Probl. Peredachi Inf.*, 1970, 6, 1, pp. 60–70.

36. Nikanorov A. A. One property of Hadamard matrices. *Probl. Peredachi Inf.*, 1974, 10, 4, pp. 95–100.

37. Nikanorov A. A. and Sagalovich J. L. New linear codes for automata. in: *Proc. Colloq. Int. Conception de Maintenance des Automatismes Logiques, Toulouse, France, 27–28 Sept. 1972*, Toulouse (1972), p. 4-4.

38. Pinsker M. S. and Sagalovich Yu. L. The memory size of automata tolerant to failures and races of delay elements, in: *Proc. Int. Workshop on Applied Issues in Automata Theory* [in Russian], Vol. 2, Varna (1971), pp. 315–324.

39. Pinsker M. S. and Sagalovich Yu. L. A lower bound on the size of automata state codes, *Probl. Peredachi Inf.*, 1972, 8, No. 3, pp. 59–66.

40. Peterson W. W. and Weldon E. J., Jr., *Error-Correcting Codes* , 1972, 2nd ed., MIT Press.

41. Pradhan D. K. and Reddy S. M., Fault-tolerant asynchronous networks, *IEEE Trans. Comput.*, 1973, 22, pp. 662–669.

42. Pradhan D. K. and Reddy S. M.Construction of $(2,1)$- separating systems using linear codes, *IEEE Trans. Comput.*, 1976, 25, pp. 945–949.

43. Pradhan D. K. Fault-tolerant asynchronous networks using read only memories, *IEEE Trans. Comput.*, 1978, 27, pp. 674–679.

44. Pradhan D. K. Asynchronous state assignment with unateness properties and fault-secure design. *IEEE Trans. Comput.*, 1978, 27, pp. 396–404.

45. C. T. Retter C. T., Intersecting Goppa codes, *IEEE Trans. Inform. Theory*, 1989, 35, pp. 822-828.

46. Sagalovich Yu. L. A method of increasing the reliability of finite automata, *Probl. Peredachi Inf.* 1965, 2, pp. 27–35.

47. Sagalovich Yu. L. An upper bound on the size of automata state codes, *Probl. Peredachi Inf.*, 1973, 1, pp. 73–83.

48. Sagalovich Yu. L. *State Encoding and Reliability of Automata* [in Russian], M.: Svyaz', 1975.

49. Sagalovich Yu. L.Information theoretical methods in the theory of reliability for discrete automata, in: *Proc. 1975 IEEE–USSR Joint Workshop Inf. Theory*, IEEE Press, New York, 1976, pp. 166–173.

50. J. L. Sagalovic. Zusammenhang von störungssicherer und wettlauffreier Kodierung der Zustände eines Automaten, in: *Dynamische Prozesse in Automaten*, VEB Verlag Techn., Berlin, 1977, pp. 95–108.

51. Sagalovich Yu. L. Some differences and connections between automata state codes and error-correcting codes, in: *Coding and Transmission of Discrete Messages* [in Russian], M.: Nauka, 1976, pp. 86–94.

52. Sagalovich Yu. L. Maximal length sequences as automata state codes, *Probl. Peredachi Inf.*, 1976, 4, pp. 70–73.

53. Sagalovich Yu. L. Cascade and generalized cascade automata state codes, in: *Proc. Sixth Int. Sympos. Inf. Theory* [in Russian], Pt. 3, Moscow–Leningrad, 1976, pp. 81–81.

54. Sagalovich Yu. L. Linear codes for automata, in: *Proc. Seventh All-Union Sympos. on Redundancy in Information Systems* [in Russian], Pt. 2, Leningrad, 1977, pp. 132–134.

55. Sagalovich Yu. L., "Cascade codes of automata states," *Probl. Peredachi Inf.*, **14**, No. 2, pp. 77–85 (1978).

56. Sagalovich Yu. L. Existence bounds for linear and nonlinear codes of automata states, in: *Proc. Eighth All-Union Conf. on Coding and Information Transmission Theory* [in Russian], Pt. 2, Coding Theory, M.: 1981, pp. 158–163.

57. Sagalovich Yu. L. Totally separating systems, *Probl. Peredachi Inf.*, 1982, 2, pp. 75–82.

58. Sagalovich Yu. L. New upper bounds on cardinality of separating systems, *Probl. Peredachi Inf.*, 1993, 2, pp. 109–111.

59. J. L. Sagalovic, "Lösung von Kodierungs- und Dekodierungsaufgaben mittels logischer Gleichungen," in: *Boolesche Gleichungen*, VEB Verlag Techn., Berlin (1984), pp. 160–174.

60. Sagalovich Yu. L., "An Improvement of the Parameter Estimates and Reduction of the Procedure of Constructing Efficient Diagnostic Polynomials," *Probl. Peredachi Inf.*, 1996, 2, pp. 77–88.

61. V. V. Sapozhnikov and Vl. V. Sapozhnikov, *Synthesis Methods for Reliable Automata* [in Russian], L.: Energiya, 1980.

62. N. J. A. Sloane, *Covering Arrays and Intersecting codes*, Preprint, 1992.

63. Solomennikov V. Yu. and Sagalovich Yu. L.. "On Linear Hash Codes",*Probl. Peredachi Inf.*, 1998, 4, pp. 13–22.

64. Tohma Y., Ohyama Y., and Sacai R., Realization of failsafe sequential machine by using $m$-out-of $n$ code, *IEEE Trans. Comput.*, 1971, pp. 1270–1275.

65. Tsfasman M. A., Vladuts S. G., and Zink T. Modular curves, Shimura curves, and Goppa codes better than the Varshamov–Gilbert bound, *Math. Nachr.*, 1982, V, pp. 21–28.

66. Unger S. H. *Asynchronous Sequential Switching Circuits*, Wiley, New York, 1969.

67. Vladuts S. G., Katsman G. L. and Tsfasman M. A. Modular curves and codes with polynomial construction complexity, *Probl. Peredachi Inf.*, 1984, 1, pp. 47–55.

68. Yakubajtis E. A. *Asynchronous Logical Automata* [in Russian], 1966, Zinatne, Riga.

69. Yakubajtis E. A. A generalized asynchronous model of finite automata, *Avtomat. Vychisl. Tekh.*, 1969, 3, pp. 1–5.

70. Zakrevzkii A. D. A synthesis method of functionally stable elements, *Dokl. Akad. Nauk SSSR*, 1959, **129**, 4, pp. 729–731.