Mathematical Theory of Redundancy Based on Formal Neuron Model

Archil Elizbarashvili
French Georgian Institute in informatics,
Faculty of Exact and Natural Sciences
Ivane Javakhishvili Tbilisi State University
University street 2, 0143, Tbilisi
GEORGIA
atchico_eliz@yahoo.com

Oleg Namicheischvili
French Georgian Institute in Informatics
Faculty of Exact and Natural Sciences
Ivane Javakhishvili Tbilisi State University
University street 2, 0143, Tbilisi
GEORGIA
oleg_namicheishvili@hotmail.com

Abstract: - Objective of the article is creation of the new methods for the binary information channels redundancy theory based on formal neuron model and development of statistic modeling algorithms and programs of computing systems, containing such channels, quorum type restoring organs reliability characteristics. Complexity of solution of such tasks is connected with the lack of the error probability evaluation of decision organ. Additional consideration and study is needed for asymptotic evaluation of error probability on the exit of threshold elements at increase of the redundancy binary channels number. It is also necessary to determine the methods ensuring for optimal weights of the threshold elements and their comparative analysis. It is probable that the optimal weight may depend on the probability of the respective channel refusal. But this latter may be only evaluated by the error frequency independent tests. Proceeding from the mentioned, it is necessary to determine statistic distribution of the input channels weights. If the majority organ shall be considered as a particular case of the threshold element, there arises a problem connected with the definition of dependence of the signal error probability received on the output of the majority organ on the input reserve channels number. It is also interesting to describe capabilities of the restoring organ in the terms of image recognition theory and determination of such threshold that would perform decision taking function with minimal risk.

Key-Words: - threshold redundancy, formal neuron, binary signal, decision, optimal weight, distribution of probabilities, method of optimization.

1 Introduction

Security of binary information channels using methods of introduction of redundancy is the major aspect of a problem of reliability. Growing interest in this problem during recent years is a result of attempts to prevent breakdown of the complex systems used in aircraft, space and power engineering and transport. Consequently development and perfection of methods of introduction of redundancy for support of working capacity of computer facilities controlling complex objects has stimulated.

Methods of introduction of redundancy known to the present moment in computers are grounded on application of
- Bodies of voting and the devices using a statistical decision theory;
- Circuits of relay type with the special architecture;
- Coding theories;
- Circuits with difficult internal structure of connections;
- Logical stabilization.

Nevertheless less the problem of rise of reliability of logical structures on the basis of model formal neuron (a threshold unit) still remains less studied and actual. In particular, question of probability of an error on an output of a threshold unit and its asymptotic behavior at increase in number of the backup channels requires further consideration. It is extremely necessary to reveal various approaches to a question of optimal assignment of weights for threshold solving (restoring) body and to carry out their comparative analysis. As optimal weight, apparently, depends on values of probabilities of errors of appropriate channels, and these probabilities are estimated as a result of independent trials by a particular of appearance of an error it is necessary to install laws of allocation of statistical characteristics of weights. If the majority restoring body is treated as a special case of threshold element the problem of reception of direct dependence of probability of an error on an output of a majority unit from number of the backup channels at $n \to \infty$ becomes inevitable.
2 Problem Formulation

Let a binary signal with the code +1 or −1 be received on the \( n \)-homogeneous information channels \( B_1, B_2, \ldots, B_n \). Due to the probable failure of the channels the quantity of variable \( x \) at the output is formulated as a set of \( x_1, x_2, \ldots, x_n \). or there is \( n \) probability of realization of variable \( x \). Clearly, each \( x_i \) \((i = 1, n)\), in its turn, is a binary variable getting the meanings +1 or −1. This redundant information \( n \) (version of the variable \( x \)) will be given to the decision receiving or restoring element (organ).

Decision receiving element is a device defining the decision, or the value of the outputting signal \( y \) on the basis of the signals \( x_1, x_2, \ldots, x_n \) given to the input. In other words, decision receiving element is a switching circuit realizing the binary function \( y = f(x_1, x_2, \ldots, x_n) \) of the binary argument \( x_1, x_2, \ldots, x_n \). If probabilities \( q_1, q_2, \ldots, q_n \) of binary channels \( B_1, B_2, \ldots, B_n \) are different, then to each information input will be ascribed its weight \( a_i \) \((i = 1, n)\), where \( a_i \) is a real number \((-\infty < a_i < \infty)\). In that case, on the basis of the weighed input signals the decision \( y \) is received using the relation \( y = \text{sgn} \left( \sum_{i=1}^{n} a_i x_i - \Theta \right) \), where \( \Theta \) is the so-called threshold of the element (quorum). Therefore, the elements working on this principle are called the threshold elements. The latter relation can also be written as \( y = \text{sgn} \left( \sum_{i=1}^{n+1} a_i x_i \right) \) if we formally assume that \( \Theta = a_{n+1} \), and \( x_{n+1} = 1 \), what means that the information channel \( B_{n+1} \) always generate the signal \( x_{n+1} = 1 \) for an arbitrary meaning of the input signal \( x \). Accordingly

\[
\text{sgn} \ z = \begin{cases} 
-1, & \text{for } z < 0 \\
0, & \text{for } z = 0 \\
+1, & \text{for } z > 0 
\end{cases}
\]

3 Problem Solution

Let us agree to consider the given problem as that of classification when the input signal \( x \) is to be ascribed to the class \( \Omega_1 \) or \( \Omega_2 \) on the basis of the \( x_1, x_2, \ldots, x_{n+1} \) versions of this signal. Number of these versions is \( n+1 \).

According to such an approach the restoring signal \( x \) must be treated as a random quantity \( X \) with realizations \( x \), characteristic of which is the \( n+1 \) random quantity \( X_1, X_2, \ldots, X_{n+1} \). Expediently the set of these quantities are to be considered a random vector. In other words, regulated combination of the column \( n+1 \) must be considered:

\[
\vec{x} = \left( X_1, X_2, \ldots, X_{n+1} \right)
\]

where \( < \ldots > \) is a symbol of transposition. Thus, each element \( X_i \) \((i = 1, n+1)\) represents a component of the random vector. Realization of the random vector \( \vec{X} \) can be written as a vector of observation:

\[
\bar{x} = (x_1, x_2, \ldots, x_{n+1})
\]

where the components of the vector \( \bar{x} \) are the realizations \( x_1, x_2, \ldots, x_{n+1} \) of the random quantities \( X_1, X_2, \ldots, X_{n+1} \).

Vector \( \bar{x} \) with the components \( X_1, X_2, \ldots, X_{n+1} \) can be described by the common function of distribution

\[
f(x) = \text{Prob}\left( X_1 = x_1, X_2 = x_2, \ldots, X_{n+1} = x_{n+1} \right),
\]

where the probable meanings of \( x_i \) \((i = 1, n)\) are −1 or +1, and \( x_{n+1} = 1 \).

It is easy to see that for the class \( \Omega_1 \) the vector \( \bar{x} \) has the following distribution

\[
f_1(\bar{x}) = \prod_{i=1}^{n+1} q_i^{1-x_i} (1-q_i)^{x_i},
\]

and for the class \( \Omega_2 \)

\[
f_2(\bar{x}) = \prod_{i=1}^{n+1} q_i^{x_i+1} (1-q_i)^{1-x_i}.
\]

Given relations are just if \( X_1, X_2, \ldots, X_{n+1} \) components of the vector \( \bar{x} \) are independent.

Mathematical expectation \( \mu = \text{M}[X] \) for any arbitrary component \( X_i \) can be obtained by particular distributions of the quantities \( X \). Combination of the aforesaid mathematical expectation \( n+1 \) can be written in the form of the mean value vector \( \vec{\mu} \):

\[
\vec{\mu} = \text{M}[\bar{x}] = \left( \mu_1, \mu_2, \ldots, \mu_{n+1} \right).
\]

Namely, the centre of distribution of probabilities of the vector \( \bar{x} \) for the class \( \Omega_1 \) can be defined by the vector

\[
\vec{\mu}_1 = \left( \mu_1, \mu_2, \ldots, \mu_n, \mu_{n+1} \right)
\]
and for the class $\mathcal{O}_2$ by

$$\tilde{\mu}_2 = \{\mu_{21}, \mu_{22}, \ldots, \mu_{2n}, \mu_{2(n+1)}\}$$

where

$$\mu_{2i} = \begin{cases} 1 - 2q_i, & i = 1, n + 1 \\ 2q_i - 1, & \text{otherwise} \end{cases}.$$  

(5)

Particular distributions of $X_i (i = 1, n + 1)$ define the dispersions $\sigma_i^2$ of these random quantities and the combined distribution of $X_i$ and $X_j$ components define their covariance $\sigma_{ij}$:

$$\Sigma_{ij} = M \left[ (X_i - \mu_i)(X_j - \mu_j) \right] = \begin{cases} \sigma_i^2, & \text{for } i = j \\ \sigma_{ij}, & \text{for } i \neq j \end{cases}.$$  

(6)

It must be noted that the combination of the dispersions $\sigma_i^2 = \sigma_j^2$ and the covariances $\sigma_{ij} = \sigma_{ji}$ make the covariance matrix generalizing the notion of one-dimensional random quantity dispersion

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \sigma_{n3} & \cdots & \sigma_{n(n+1)} \end{bmatrix}.$$  

As we consider the case, when the components $X_1, X_2, \ldots, X_{n+1}$ of the vector $\tilde{x}$ are not interdependent, hence $\sigma_{ij} = 0$ for every $i \neq j$ and it represents a diagonal matrix $\Sigma$:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \sigma_{n(n+1)}^2 \end{bmatrix}.$$  

(7)

It can easily be seen that the dispersions $\sigma_i^2$ are equal for both classes and is defined by the expression

$$\sigma_i^2 = \sigma_j^2 = \sigma^2 = 4q_i(1 - q_i).$$

Consequently, the covariance matrices are equal

$$\Sigma_1 = \Sigma_2 = \Sigma.$$  

(8)

So, studying the threshold element we assume that the parameters $\hat{\mu}_1, \hat{\mu}_2$ and $\Sigma$ are known and according to its working principle the threshold organ is considered to be a linear combination of observations $z = a_1X_1 + a_2X_2 + \cdots + a_{n+1}X_{n+1}$.

This expression is called a linear discriminator. The vector of observations $\tilde{x}$ is considered to belong to the class $\mathcal{O}_1$ for $z > 0$ and to the $\mathcal{O}_2$ for $z < 0$, and for $z = 0$ the decision is not received.

Let us introduce to consider a random quantity $Z$ defining it by the relation

$$Z = \sum_{i=1}^{n+1} a_iX_i,$$  

(9)

or

$$Z = \sum_{i=1}^{n+1} Z_i,$$  

(10)

where

$$Z_i = a_iX_i.$$  

(11)

If the observation $\tilde{x}$ for the class $\mathcal{O}_1$ is realized, then the sum (9) has the distribution

$$F_1(z) = \sum_{i=1}^{n+1} f_1(z_i),$$  

(12)

where $* -$ is a symbol of convolution, and

$$f_1(z_i) = q_i^{2a_j} (1 - q_i)^{2a_j}.$$  

(13)

Similarly, when the observation $\tilde{x} = (x_1, x_2, \ldots, x_{n+1})$ belong to the class $\mathcal{O}_2$, then distribution of probabilities of the random quantity $Z$ is defined by the formula

$$F_2(z) = \sum_{i=1}^{n+1} f_2(z_i),$$  

(15)

where

$$f_2(z_i) = q_i^{2a_j} (1 - q_i)^{2a_j} z_i.$$  

(16)

and the probable values of $z_i$ are $+a_j$ and $-a_j$. If the observation $\tilde{x}$ is of the class $\mathcal{O}_1$, then mathematical expectation of the random quantity $Z$ is defined by the formula

$$m_1 = \sum_{i=1}^{n+1} a_i\mu_i = \sum_{i=1}^{n+1} a_i(1 - 2q_i).$$  

(14)

Similarly, when the observation $\tilde{x} = (x_1, x_2, \ldots, x_{n+1})$ belong to the class $\mathcal{O}_2$, then distribution of probabilities of the random quantity $Z$ is defined by the formula

$$m_2 = \sum_{i=1}^{n+1} a_i\mu_i = \sum_{i=1}^{n+1} a_i(1 - 2q_i).$$  

(17)

Comparing the formulae (14) and (17) we can be easily conclude that $m_i = m_2$, and analysing the formulae (13) and (16) we derive

$$f_2(z_i) = f_1(-z_i).$$  

(19)

Consequently,

$$F_1(z) = F_1(-z).$$  

(20)

Dispersion $\sigma_z^2$ of the random quantity $Z$ is equal in both cases and is defined by the formulae

$$\sigma_z^2 = \sum_{i=1}^{n+1} \sum_{j=1}^{n+1} a_i a_j \sigma_{ij}.$$  

(21)

Using (7) for $\sigma_{ij} = 0$ the formulae (21) gives

$$\sigma_z^2 = \sum_{i=1}^{n+1} 4a_i^2 q_i(1 - q_i)$$  

(22)
According to the heuristic points of view the weights $a_1, a_2, \ldots, a_{n+1}$ should be chosen so that the values of the mathematical expectations $m_1$ and $m_2$ be different as far as possible, and the dispersion $\sigma^2$ be minimal. It suffices to choose the (Mahalanobis') generalised distance [2] for a target function
\[ \rho = \frac{(m_1 - m_2)^2}{\sigma^2}. \] (23)
In our case it has the form:
\[ \rho = \frac{\left( \sum_{i=1}^{n+1} a_i (\mu_{1i} - \mu_{2i}) \right)^2}{\sum_{j=1}^{n+1} \sum_{i=1}^{n+1} \sigma^2_{ij} a_j}. \] (24)
The weights $a_i$ ($i = 1, n + 1$) giving the maximum to this expression satisfy the following system
\[ \frac{\partial \rho}{\partial a_i} = 0, \quad i = 1, n + 1. \] (25)

Using (24) we arrive at
\[ \frac{\mu_{1s} - \mu_{2s}}{\sum_{j=1}^{n+1} \sum_{i=1}^{n+1} \sigma^2_{ij} a_j} = \sum_{i=1}^{n+1} a_i (\mu_{1i} - \mu_{2i}). \] (26)
Each vector $(a_1, a_2, \ldots, a_{n+1})'$ is the solution of the system of equations (26) if it satisfies the system
\[ \sum_{j=1}^{n+1} \sigma^2_{ij} a_j = k (\mu_{1s} - \mu_{2s}), \quad s = 1, n + 1, \] (27)
where $k$ is an arbitrary (free) constant.

Taking into account that in (23) $\sigma^2_{ij} = 0$ for $i \neq j$ and $\sigma^2_{ii} = \sigma^2_i$ we get
\[ a_i \sigma^2_i = k (\mu_{1s} - \mu_{2s}), \quad s = 1, n + 1, \] from where
\[ a_s = \frac{k (\mu_{1s} - \mu_{2s})}{\sigma^2_i}, \quad s = 1, n + 1. \] (28)

Using (5) and (7) finally we get
\[ a_i = \frac{k (1 - 2q_i)}{2q_i (1 - q_i)}, \quad \frac{1}{n + 1}, \] (29)
If the existence of positive weights is desirable for $q_i < 1/2$ and the negative weights for $q_i > 1/2$, then the constant $k$ must satisfy the condition $0 < k < \infty$.

For such weights the distance $\rho$ equals the absolute value of the difference of the characteristics $m_1$ and $m_2$.

\[ \rho = |m_1 - m_2| = \sum_{i=1}^{n+1} \frac{(1 - 2q_i)^2}{q_i (1 - q_i)} \] (30)

where, as it was already mentioned, $m_i$ ($i = 1, 2$) is a mathematical expectation of the sum $Z$ for the class $\Omega_i$. Thus, if we choose the weights $a_i$ according to (29) then there holds a theorem
\[ \frac{\sigma^2}{|m_1 - m_2|} = 1. \] (31)

Formula (30) shows that if probabilities of errors at the inputs do not equal 1/2, then increasing the number of threshold element inputs $n$ the generalized distance $\rho$ is monotonously increased and the probability of signal restoration error is reduced. In particular, generalized distance $\rho$ for majority element equals
\[ \rho = \frac{(1 - 2q)^2}{q (1 - q)} n \] (32)
where $q_1 = q_2 = \cdots = q_n = \frac{1}{2}$ and $q_{n+1} = 1/2$.

Family of the dependences $\rho(q)$ for a number of meanings $n$ is shown in Fig.2.

![Fig.2. Dependence of (Mahalanobis') generalized distance on the probability of error of majority element inputs for some values of the number of the inputs.\rho](image)

**2.1 Probability of error.**
Let us consider the Fig.3 to define the probability of error of the threshold element. There are shown the two distributions ($F_1(z)$ and $F_2(z)$) of the random quantity $Z$ for the origin $z = 0$.

![Fig.3. Discontinuous (discrete) distributions $F_1(z)$ and $F_2(z)$ of the random sum $Z$ for the classes $\Omega_1$ and $\Omega_2$.](image)
If \( x=1 \), i.e. if the vector of observations \( \bar{x} \) belongs to \( \Omega_1 \) but \( z<0 \), then the decision \( y=-1 \) is mistaken and its probability is defined by the expression:

\[
Q_1 = \sum_{z<0} F_1(z) \tag{33}
\]

If \( x=-1 \), i.e. if the vector of observations \( \bar{x} \) belongs to \( \Omega_1 \), but realization of the random quantity \( Z \) is more than zero \( (z>0) \), then the decision is mistaken and its probability is defined by the expression:

\[
Q_2 = \sum_{z>0} F_2(z) = \sum_{z<0} F_1(-z) = \sum_{w<0} F_1(w) = Q_1 \tag{34}
\]

Probability \( Q \) of the signal restoration error of threshold decision organ can be defined by the formula of the complete probability

\[
Q = q_{n+1} Q_1 + (1 - q_{n+1}) Q_2 = Q_1 + Q_2 \tag{35}
\]

where \( q_{n+1} \) is a priori probability of identification the class \( \Omega_1 \), i.e. a priori probability that the signal \( x \) has the meaning \( x=+1 \). Thus

\[
Q = \sum_{z<0} F_1(z) = \sum_{z<0} f_{i_1}(z_i), \tag{35}'
\]

where the function \( f_{i_1}(z_i) \) is defined by (13) and there is to be introduced only that member which corresponds to the negative meaning of the variable \( z \). The total number of discrete meanings of the variable \( z \) is \( 2^{n+1} \) as \( z = \tilde{a}_1 + \tilde{a}_2 + \cdots + \tilde{a}_n + \tilde{a}_{n+1} \), where \( \tilde{a}_i \) equals \( +a_i \) or \( -a_i \).

According to the formula (35)' each discrete quantity of the variable \( z \) is satisfied by the summand \( Q_j \) (\( j=1, n+1 \)) equal to the product of probabilities:

\[
Q_j = F_j(z) = \prod_{i=1}^{n+1} f_i(z_i) = \tilde{q}_1 \cdot \tilde{q}_2 \cdot \cdots \tilde{q}_n \cdot \tilde{q}_{n+1}
\]

where

\[
\tilde{q}_k = \begin{cases} 
q_k, & \text{for } z_k = -a_k \\
1 - q_k, & \text{for } z_k = a_k
\end{cases}
\]

To derive the existence of probability \( Q \) it suffices to sum up the summands answering the negative meanings of the variable \( z \), i.e.

\[
Q = \sum_{z<0} Q_j = \sum_{z<0} \tilde{q}_1 \cdot \tilde{q}_2 \cdot \cdots \tilde{q}_n \cdot \tilde{q}_{n+1}. \tag{35}''
\]

Evidently, the exact algorithm obtained this way coincides the basic result of [3].

### 4 Conclusion

Method of optimization of (Mahalanobis') generalized distance has been applied for the first time to define the weights of the threshold element inputs. Quite a new result (29) is obtained. Determination of probability of error of the threshold element with the approaches of pattern recognition theory led us to the algorithm (35)" with the weights estimated by optimization of Mahalanobis' distance.

### References:

