Buletinul Ştiinţific al Universităţii din Baia Mare
Seria B, Matematică-Informatică, vol.XII (1996),139-154

Dedicated to the 35th anniversary of the University of Baia Mare

Rational Algorithm of Recognition

by
Petru Avram PETRIŞOR

1. INTRODUCTION. Information is an objective property of the material processes and the notion of information is often identified with the notion of quantity of information although within the limites of a formal model the quantity of information could be mathematically defined as Shannon ([10]) , did it 1948 by using the notion of entropy. This way of studying the quantity of information is unsubstantial because it doesn't show us the important informational characteristics of the studied phenomenon, that is those properties of the physical world that are opposed to a certain extent to the energetical or mass characteristics.

In the analysis of the notion of quantity of information there appear no proportions with a statistic character of the entropy -type but another type of characteristics that refer to the contents of information and not to the statistic properties of information. From the statistic point of view the study of information implies the possibility to draw out of a given communication a certain quantity of information, but it can't be specified what kind of information can be drawn out of the mentioned communication. On the other hand the information is strongly bound to its transmission and depending on the hypothesis of the unity of the material world, this transmission can't be surpass the limits of reality, that is information doesn't exist (the information) represents a special relation between material processes and so it is a characteristic of the substance (matter) that can't be identified with information.

The concept of information appeared slowly but spectacular achieving a synthesis of preoccupations apparently unbound to each other.

It was followed by the generation of creative ideas in other independent domains.

2. CERTAINTY AND UNCERTAINTY. The generation of an event brings about the more information the more it is more unexpected and this characteristic of the event can be measured if its probability of happening is known, that is, the information brought about by this event is so much bigger as this probability is smaller.

Be it $\mathcal{A}$ a set of events echiprobabile that are the results of an experiment having many possible results. If these results are equally probable then by effectuation of the experiment the information produced depends on the fact that there were more results having equal probabilities and so there don't exist N results having the $\frac{1}{N}$ probability.

Consequently, the quantity of information, when one of these results is obtained, is a decreasing function of $\frac{1}{N}$ or an increasing function of N noted with f(N) and thus we can write: f(N)=a log b, where a>0 and b>0 are constantly ([1]). If we choose b=2, a=1 and $p=\frac{1}{N}$ we obtain: $f(N)=\log_2 p$ where p is the probability of appearance of an event that contains a result. In case that the event generates a finite number of events noted with $A_1, A_2, \ldots A_n$ having probabilities in $p_1, p_2, \ldots p_n$ it can be deduced that, if $p_1=1$, $p_i=0$ (i=2,3, ...n) then the experiment doesn't bring anything new, producing a certainty.

The experiment that generates the events $A_1, A_2, \ldots A_n$ having the probabilities $p_1, p_2, \ldots p_n$ contains a certain uncertainty concerning the result of the experiment taking into account and the degree of uncertainty is measured by means of the expression:

$$(1)\ H(A) = -\sum_{i=1}^{n} p_i \log_2 p_i$$

([10]) that is called informational entropy and it has a statistic character. Admitting that the experiment will take place it can be deduced that $p_1+p_2+\ldots+p_n=1$.

To obtain the extremities of the function H that has the variables $p_1, p_2, \ldots p_n$ we use Lagrange's method of the multipliers ([4]) and in these case may it be considered the following function:

$$(2)\ F(p_1, p_2, \ldots p_n)=H(p_1, p_2, \ldots p_n)+\lambda(p_1+p_2+\ldots+p_n)$$

The extremes of function F can be obtained by solving the system:

$$(3)\quad \begin{cases} \dfrac{\partial}{\partial p_1} F(p_1, p_2, \ldots p_n) = 0 \\ \dfrac{\partial}{\partial p_2} F(p_1, p_2, \ldots p_n) = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dfrac{\partial}{\partial p_n} F(p_1, p_2, \ldots p_n) = 0 \\ p_1 + p_2 + \ldots + p_n = 1 \\ p_1 > 0, p_2 > 0, \ldots, p_n > 0 \end{cases}$$

This system is equivalent with the system:

$$(4)\quad \begin{cases} -\dfrac{1}{\ln 2}(1 + \ln p_1) = -\lambda \\ -\dfrac{1}{\ln 2}(1 + \ln p_2) = -\lambda \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ -\dfrac{1}{\ln 2}(1 + \ln p_n) = -\lambda \\ P_1 + p_2 + \ldots + p_n = 1 \\ p_1 > 0, p_2 > 0, \ldots, p_n > 0 \end{cases}$$

This system has the solution $p_1=p_2=\ldots=p_n=\dfrac{1}{n}$ and taking into account that the second differential of the function H calculated in the point $\left(\dfrac{1}{n}, \dfrac{1}{n}, \ldots \dfrac{1}{n}\right)$ is negative that is $d^2H<0$ it can be deduced that H is maximum if all the probabilities p_j $(j=1,2,\ldots n)$ are equal among them, so if $H_{max}=H\left(\dfrac{1}{n}, \dfrac{1}{n}, \ldots \dfrac{1}{n}\right)$ and H(A) is the informational entropy of the event A then the following relation takes place:

(5) $H(A)=H\leq H_{max}=-\log_2 p$

The informational entropy measures the quantity of information J that is contained in an experiment and it can be determined by using the relation:

(6) $J= H_{max} - H$

If H_{max} is a constant size, then, the more the event A has bigger chances of achieving the smaller is the quantity of information that its achieving determines.

There is a certain limit of growing of the quantity of information, over which, when, passing over, we come to the producing of some perturbation, consequently there is a maximum quantity of necessary information in relation with which the informational entropy has the value zero.

If H_t is the informational energy at the moment "t" and H_u is the informational energy at the moment "u" then $J_{ut}=H_u-H_t$ represents a quantity of information existing between the two moments and taking into account that $H_t=H_u-J_{ut}$ it can be deduced that the informational entropy decreases at the same time with the growing of the quantity of information existing in a experiment ([9]|).

With the social systems by introducing the positive subunitary number $r=J:H_{max}$ and taking into account that the lack of sufficient information causes (brings about) disorganisation of the social system, it can be deduced that the stability of a system grows inversely proportional to the informational entropy an so the number "r" characterises the degree of organisation of a social system.

On the other hand the social systems are characterised by the growing of complexity and consequently of the maximum informational entropy that imply the decreasing of the quantity of information existing in a system.

Consequently the growing of the complexity of the information of a social system is achieved by the growing of the quantity of information and this can be obtained by especially developping the educational system and the sciences in order to establish a stability (an equilibrium) between complexity and organisation.

The unit of measure of the information is called binary unity or bit and is the smallest possible unity of quantity of information.

If we take the bit as unity of measure of the quantity of existing information in each symbol of the alphabet of the French language and taking into account that it has 26 letters and supposing that each letter is used with the same frequency we obtain:

$$(7)\ H = -\sum_{i=1}^{25} \frac{1}{26} \log_2 \frac{1}{26} = \log_2 26$$

Consequently, the quantity of information that a letter of the alphabet has involved in it, is of 5,506 bits and so there can be calculated the quantity of information of different messages.

3. INFORMATIONAL ENERGY. The development of certain human activities isn't possible without a change of informations. If the energy and the mass exist always in the shape of certain material objects, the information exists in the shape of a message about an experiment and in the case when there exist more messages about the same experiment we'll choose that message that is more useful

and complete and in order to make this choice, it is necessary to dispose of a measure of information that can be obtained by doing an experiment and this fact eliminates a certain indetermination. Consequently, we can say that the information replaces an indetermination that it can be measured, we could have the possibility to measure the information.

The utility of information is closely tied to the way of measuring the quantity of information and one of the first answers was given by R.A.Fischer ([5]) and afterwards by R.V.H.Hartley ([6]) but the final solution was given by C.Shannon ([10]).

The concept of the utility of information appeared for the first time in 1730 in G.Cramer's works (1704-1752) and it was resumed by D.Bernoully (1700-1782) when the theory of probabilities was rigurously proved scientifically, but the study of the utility of information from the mathematical point of view was achieved by T. von Newman (1903-1957) and O.Morgenstern (1902-1977) who formulated the utility as a measure of preferance for a one result of the experiment or another one.

The global information of a system S with the positions $s_1, s_2, \ldots, s_n$ having the weight $p_1, p_2, \ldots p_n$ is given by the informational energy and that can be calculated by using the expression:

$$(8)\ E_s = p_1^2 + p_2^2 + \ldots + p_n^2$$

where $p_1+p_2+\ldots+p_n=1$ ([3], [9]).

We notice that at the same time with the growing of the disorganisation of the system the decreasing of the informational energy takes place.

Really, be it s_i an s_j two positions of the system S having the weight in p_i respectively in p_j so that $p_j<p_i$. After the changing of the system the weight in p_i' becomes $p_i'=p_i-x$ and the weight in p_j becomes $p_j'=p_j+x$ because we must have $p_i+p_j=p_i'+p_j'$ and consequently the variable point of the expression E_s has the following form:

$$(9)\left(p_i'\right)^2 + \left(p_j'\right)^2 = p_i^2 + \left(p_j'\right)^2 + 2x^2 - 2x\left(p_i - p_j\right)$$

If $p_i'-p_j'=p_i-p_j-2x$ grows, that is $x<0$ then the following inequality takes place:

$$(10)\ \left(p_i'\right)^2 + \left(p_j'\right)^2 > p_i^2 + p_j^2$$

and so the informational energy grows. If $0<x<p_i-p_j$ then the following inequality takes place:

(11) $p_i^2 + p_j^2 < (p_i^{'})^2 + (p_j^{'})^2$

and consequently the informational energy decreases at the same time with the decreasing of the expression p_i-p_j remaining in the same relationship p_i>p_j.

Concluding, the informational energy grows with the organisation of the system and decreases at the same time with the information or disorganisation of the system.

In the pedagogical experiment it is considered that the most adequate from for the achieving of performance are the texts with multiple choice. In a multiple choice text an information is received and an answer is chosen to the alternatives that are offered. One of the answers is correct and the others are wrong representing the most frequent mistakes.

To obtain the most favourable number of possible answers, in a multiple choice text, so that the maximum of information may be obtained, the informational energy is applied to Chernoff 's model ([2]).

Be it p_i the absolute frequency of the subjects that choose correctly the answer to the question i,q_i the absolute frequency of the subjects that choose the wrong answer to the question i and w_i the absolute frequency of the subjects that omit the answer to the question is consequently for each i, the following equality takes place:

(12) $p_i+q_i+w_i=1$

and if y_i is absolute frequency of the subjects that know the correct answer to the question i and they don't choose it at random, then Chernoff 's relation for a text with m possible answers referring to the question i is:

(13) $p_i = y_i + \frac{1}{m}(1 - w_i - y_i), i = 1,2, \ldots, k.$

Taking into account the equality $p_i=y_i+(p_i+y_i)$ and the expression:

(14) $E_i = y_i^2 + (p_i - y_i)^2 + q_i^2 + w_i^2$

of the informational energy of the question of rank i it may be deduced that E_i has an extreme in relation to the variable y_i if y_i is the solution of the system :

(15) $\frac{\partial E_1}{\partial y_1} = 0, \frac{\partial E_2}{\partial y_2} = 0, \ldots, \frac{\partial E_k}{\partial y_k} = 0$

that is equivalent to the system :

(16) $2y_i-2(p_i-y_i)=0, i=1,2,\ldots,k$

and of which we can obtain the value $y_i = \frac{p_i}{2}$, $(i = 1,2, \ldots, k)$

Taking into account the relation (13) we obtain :

$$(17) \quad m = \frac{2 - 2w_i - p_i}{p_i}, i = 1,2, \ldots, k.$$

The relation (17) determines the most favourable number of possible answer to a test with a multiple choice so that the maximum quantity if information may be obtained.

4. ALGORITHM OF RECOGNITION. An old notion used already by Euclid, heaving its name taken from the Arab man of science Al Horezmi (680-846), a notion often met with in mathematics and not only in mathematics, a notion very important for our days, is the notion of algorithm.

An important problem put in the theory of algorithms is to find the classes of problems that can be solved algorithmically and to build up algorithms for the solving of the classes of problems. Taking into account the importance of the algorithms and the fact that by using them we may obtain a rigorous study of the problem analysed we deduce hence that the discovery and mastering of the algorithms is a problem of psychology and pedagogy.

The algorithms of recognition have a special importance because by using them we can establish the class of problems to which a problem belongs.

The algebra of propositions play an essential part in the theory of algorithms because the algorithm is a finite row (line) of instructions that are achieved in a given order and that has a finality the solving of a problem of given class.

By the symbol p we understand a proposition if p is a true or false statement but not true or false at the same time, and by A(p) we understand that the sentence p is true and F(p) represents the fact that p is false. The sentence "whatever the sentence p may be the A(p) or F(p)" takes place. This is called the principle of the third term excluded, and the sentence "doesn't exist any sentence p so that A(p) and F(p)" is called the principle of contradiction and the ensemble of the two principles constites the principle of bivalency.

From the principle of bivalency it doesn't result that for each sentence p there exists a method by which it can be established that p is true or that p is false. There are mathematical sentences about which we don't know even up in our days if they are true or false. For instance the problem of the problem of the independence of the hypothesis of the continuum up to the year 1963 was such a sentence and after this date this sentence is true.

The notions of "true" of "false" aren't put into evidence by the logic but by the special sciences, these notions entering into the logic as postulated entities.

In other words there is admitted the initial existence of two symbols separated from the others, the one to designate the truth, the other to designate the false marked with 1 respectively 0.

If $\mathcal{P}$ is a set of sentences then we note with $\mathcal{P}_0 \subset \mathcal{P}$ the subset of false sentences and with $\mathcal{P}_1 \subset \mathcal{P}$ the subset of true sentence.

Consequently:

(19) $P=P_0 \cup P_1,\ P_0 \cap P_1=\varnothing$.

We notice that by means of some linking words from a given multitude of sentences new sentences may be obtained.

This linking words are called logical connectives and they correspond in a certain sense to the coordinative or subordinative conjunctions in grammar. These connectives are the following:

a) The connective "non" is applied to only one sentence p and forms the sentence "non p" called the negation of the sentence p;

b) The connective " and" applies to two sentences p_1, p_2 and forms the sentence $p_1 \wedge p_2$ called conjunction of the sentences p_1 and p_2;

c) The connective "or" is applied to two sentences p_1, p_2 and it forms a new sentence $p_1 \vee p_2$ called disjunction of the sentences p_1 and p_2;

d) The connective "if... then" is applied to two sentences p_1, p_2 forming the sentence "if p_1, then p_2 " called the implication of the sentences p_2 by p_1 (p_1 implies p_2). This connective is called also the conditional of p_1 and p_2 keeping the denomination of implication for the case when the conditional is true;

e) The connective "...if and only if..." applies to two sentences p_1, p_2 forming a new sentence "p_1 if and only if p_2 " called the equivalence of the two sentences. The set $B_2=\{0,1\}$ is considered where 0 symbolises the notion of false and 1 the notion of true and be it the function:

$$v: P \to B_2,\ v(p) = \begin{cases} 1 \text{ if } p \text{ is true} \\ 0 \text{ if } p \text{ is false} \end{cases}$$

From the principle of contradiction it can be deduced that the function "v" is uniform, that is to each proposition corresponds a single value, 0 or 1 and the principle of excluded middle established that the function v is a subjection.

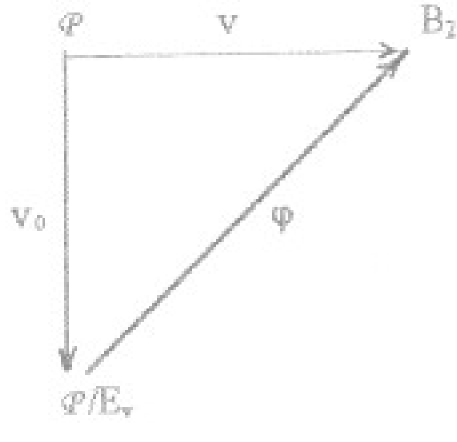
To point out the difference between this function and a certain function from mathematics, we take into account the following example:

$$f: N \rightarrow N, f(n) = \begin{cases} 0 \text{ if n is even} \\ 1 \text{ if n is uneven} \end{cases}$$

Comparing the two functions we find out that we don't dispose of a logical operation with which to calculate v(p) for a proposition p. In the moment when p is precised, it belongs to a scientific field and by the methods adequate to the respective scientific discipline we can establish if the proposition p is false or true, but the proceeding for this verification belongs to the particular scientific field and that means that isn't a logic operation defined on the set $\mathcal{P}$. The incalculability of the function v is due to the general definition of the notion of proposition.

The parity and imparity of the natural number n that makes the object of the function f may be verified by means of the operations defined on the set N: the devision of the number n to the number 2.

By means of the function v we may define on $\mathcal{P}$ a relation equivalence like this: if $p,q \in \mathcal{P}$ then p is in relation E_v with quantity if $E_v(p,q)=v(p)=v(q)$. If $\mathcal{P}/E_v$ is the factorisation of the set $\mathcal{P}$ by means of the relation E_v and $v_0:\mathcal{P}\rightarrow\mathcal{P}/E_v$ is the natural application, then we obtain the canonic decomposition of the function v given by the diagram:



and $\varphi([p]_{E_v})= v(p)$, where $[p]_{E_v}$ is the class of equivalence in connection to the relation E_v generated by the proposition p.

From the definition of the connectives we obtain the following properties:

(α) The disjunction "∨" of two propositions is only then false when both propositions are false;

(β) The conjunction "∧" of two propositions is true only when both propositions are true;

(χ) The implication "→" of two propositions is false when the first proposition is true, and the second is false;

(δ) The equivalence "↔" is true when the two propositions are either both true or are both false;

(ε) The negation "⌉" of a true proposition is a false proposition and the negation "⌉" of a false proposition is a true proposition.

On the other hand taking into account that for each propositional function ⌉, ∧, ∨, →, ↔ there exist the booleeane f_{13}, f_{28}, f_{22}, f_{25}, f_{27} so that B_2 in relation with f_{28}, f_{22} and f_{13} is a Boole algebra we deduce that the set of propositions constitute a Booleean algebra in relation to the functions of truth ⌉, ∧, ∨ ([8]).

In this case the following formulae take place:

(α_1) $(p_1 \wedge p_2) \equiv p_1 \quad (\forall)\, p_1 \in \mathcal{P}$;

(α_2) $(p_1 \wedge (p_2 \wedge p_3)) \equiv ((p_1 \wedge p_2) \wedge p_3) \quad (\forall)\, p_1, p_2, p_3 \in \mathcal{P}$;

(α_3) $(p_1 \wedge p_2) \equiv (p_2 \wedge p_1) \quad (\forall)\, p_1, p_2 \in \mathcal{P}$;

(α_4) $(p_1 \wedge (p_2 \vee p_3)) \equiv (p_1 \wedge p_2) \vee (p_1 \wedge p_3) \quad (\forall)\, p_1, p_2, p_3 \in \mathcal{P}$;

(α_5) $\rceil(p_1 \wedge p_2) \equiv ((\rceil p_1) \vee (\rceil p_2)) \quad (\forall)\, p_1, p_2 \in \mathcal{P}$;

(α_6) $(p \wedge \rceil p) \quad (\forall)\, p \in \mathcal{P}$ is a false proposition;

(α_7) If p_1 and p_2 are equivalent propositions, then the propositions, then the propositions: $(\rceil p_1) \wedge p_2$ and $p_1 \wedge (\rceil p_2)$ are false;

(α_1') $(p_1 \vee p_2) \equiv p_1 \quad (\forall)\, p_1 \in \mathcal{P}$;

(α_2') $(p_1 \vee (p_2 \vee p_3)) \equiv ((p_1 \vee p_2) \vee p_3) \quad (\forall)\, p_1, p_2, p_3 \in \mathcal{P}$;

(α_3') $(p_1 \vee p_2) \equiv (p_2 \vee p_1) \quad (\forall)\, p_1, p_2 \in \mathcal{P}$;

(α_4') $(p_1 \vee (p_2 \wedge p_3)) \equiv (p_1 \vee p_2) \wedge (p_1 \vee p_3) \quad (\forall)\, p_1, p_2, p_3 \in \mathcal{P}$;

(α_5') $\rceil(p_1 \vee p_2) \equiv ((\rceil p_1) \wedge (\rceil p_2)) \quad (\forall)\, p_1, p_2 \in \mathcal{P}$;

(α_6') $(p \vee \rceil p) \quad (\forall)\, p \in \mathcal{P}$ is a true proposition;

(α_7') If p_1 and p_2 are equivalent propositions, then the propositions, then the propositions $(\rceil p_1) \vee p_2$ and $p_1 \vee (\rceil p_2)$ are true proposition;

In these properties the sign ≡ represents the echivalence of the propositions, that is the value of truth of the two members are equal.

We'll suppose now, that in the process of learning at a certain level it is necessary to establish if the proposition $p \in \mathcal{P}$ is one of the following types:

A) The proposition has subject and predicate;

B) The proposition has subject and no predicate in the Future Tense;

C) The proposition has no subject and the predicate is expressed by a verb in the first or second person;

D) The proposition has no subject and the predicate is expressed by a verb in the Past Tense, in the third person plural;

E) The proposition has no subject and the predicate is expressed by a verb in the first or second person any tense and the predicate isn't expressed by a verb in the Past Tense, in the third person plural;

We consider, going on, the following characteristics:

(a) The proposition has subject;

(b) The proposition has predicate;

(c) The proposition has predicate expressed by a verb in the first or second person;

(d) The proposition has predicate expressed by a verb in the Past Tense, the third person plural;

Taking into account these definitions and the propositional functions $\neg$, $\wedge$, $\vee$ as the properties (α_i), (α_i') $i=1,\dots,7$ we can write:

(18) $A \equiv (a\wedge b\wedge c\wedge d)\vee(a\wedge b\wedge c\wedge(\neg d))\vee(a\wedge b\wedge(\neg c)\wedge d)\vee(a\wedge b\wedge(\neg c)\wedge(\neg d))$;

(19) $B \equiv (a\wedge(\neg b)\wedge c\wedge d)\vee(a\wedge(\neg b)\wedge c\wedge(\neg d))\vee(a\wedge(\neg b)\wedge(\neg c)\wedge d)\vee$
$(a\wedge(\neg b)\wedge(\neg c)\wedge(\neg d))$;

(20) $C \equiv ((\neg a)\wedge b\wedge c\wedge d)\vee((\neg a)\wedge b\wedge c\wedge(\neg d))\vee((\neg a)\wedge(\neg b)\wedge c\wedge d)\vee((\neg a)\wedge$
$(\neg b)\wedge c\wedge(\neg d))$;

(21) $D \equiv ((\neg a)\wedge b\wedge c\wedge d)\vee((\neg a)\wedge b\wedge(\neg c)\wedge d)\vee((\neg a)\wedge(\neg b)\wedge c\wedge d)\vee((\neg a)\wedge$
$(\neg b)\wedge(\neg c)\wedge d)$;

(22) $E \equiv ((\neg a)\wedge b\wedge(\neg c)\wedge(\neg d))\vee((\neg a)\wedge(\neg b)\wedge(\neg c)\wedge(\neg d))$;

We notice that the propositions $(\neg b)\wedge c$, $(\neg b)\wedge d$, $(\neg a)\wedge(\neg b)$, $c\wedge d$ are false and thus we have:

(23) $A \equiv (a\wedge b\wedge c\wedge(\neg d))\vee(a\wedge b\wedge(\neg c)\wedge d)\vee(a\wedge b\wedge(\neg c)\wedge(\neg d))$;

(24) $B \equiv (a\wedge(\neg b)\wedge(\neg c)\wedge(\neg d))$;

(25) $C \equiv ((\neg a)\wedge b\wedge c\wedge(\neg d))$;

(26) $D \equiv ((\neg a)\wedge b\wedge(\neg c)\wedge d)$;

(27) $E \equiv ((\neg a)\wedge b\wedge(\neg c)\wedge(\neg d))$;

Taking into account the preceding relations we have for each proposition $p \in \mathcal{P}$:

(28) $P \equiv (a \wedge b \wedge c \wedge (\neg d)) \vee (a \wedge b \wedge (\neg c) \wedge d) \vee (a \wedge b \wedge (\neg c) \wedge (\neg d)) \vee (a \wedge (\neg b) \wedge (\neg c) \wedge (\neg d)) \vee ((\neg a) \wedge b \wedge c \wedge (\neg d)) \vee ((\neg a) \wedge b \wedge (\neg c) \wedge (\neg d))$;

We'll introduce the notations:

$P_{AB} \equiv (a \wedge b \wedge c \wedge (\neg d)) \vee (a \wedge b \wedge (\neg c) \wedge d) \vee (a \wedge b \wedge (\neg c) \wedge (\neg d)) \vee (a \wedge (\neg b) \wedge (\neg c) \wedge (\neg d))$;

$P_{CDE} \equiv ((\neg a) \wedge b \wedge c \wedge (\neg d)) \vee ((\neg a) \wedge b \wedge (\neg c) \wedge d) \vee ((\neg a) \wedge b \wedge (\neg c) \wedge (\neg d))$;

$P_{DE} \equiv ((\neg a) \wedge b \wedge (\neg c) \wedge d) \vee ((\neg a) \wedge b \wedge (\neg c) \wedge (\neg d))$;

By means of these notations we may write:

(29) $P_{AB} \equiv A \vee B$;

(30) $P_{CDE} \equiv C \vee D \vee E)$

(31) $P_{DE} \equiv D \vee E$;

By the notation $a(P) \in H_M$ we understand that out of all characteristics of the proposition P the characteristic 'a' has the biggest informational entropy.

In a similar manner we'll interpret the notations we $b(P_{AB}) \in H_M$, $c(P_{CDE}) \in H_M$ and $d(P_{DE}) \in H_M$.

The characteristics a, b, c, d have different probabilities that are determined by a certain statistics of the language that is we take a big amount of propositions and we calculate the frequency of the characteristic a,b,c,d. This statistics is the more edifying, the bigger the number of the analysed proposition is. Like this, we obtain the chart:

	a	b	c	d
P	$p_0(a)$	$p_0(b)$	$p_0(c)$	$p_0(d)$
	$H_0(a)$	$H_0(b)$	$H_0(c)$	$H_0(d)$
P_{AB}	$p_1(a)$	$p_1(b)$	$p_1(c)$	$p_1(d)$
	$H_1(a)$	$H_1(b)$	$H_1(c)$	$H_1(d)$
P_{CDE}	$p_2(a)$	$p_2(b)$	$p_2(c)$	$p_2(d)$
	$H_2(a)$	$H_2(b)$	$H_2(c)$	$H_2(d)$
P_{DE}	$p_3(a)$	$p_3(b)$	$p_3(c)$	$p_3(d)$
	$H_3(a)$	$H_3(b)$	$H_3(c)$	$H_3(d)$

where:

$H_i(a)= -p_i(a) \log_2 p_i(a) - p_i(\neg a) \log_2 p_i(\neg a)$, $i=0,1,2,3$;

$H_i(b)= -p_i(b) \log_2 p_i(b) - p_i(\neg b) \log_2 p_i(\neg b)$, $i=0,1,2,3$;

$H_i(c)= -p_i(c) \log_2 p_i(c) - p_i(\neg c) \log_2 p_i(\neg c)$, $i=0,1,2,3$;

$H_i(d)= -p_i(d) \log_2 p_i(d) - p_i(\neg d) \log_2 p_i(\neg d)$, $i=0,1,2,3$;

We suppose that the proposition $p \in \mathcal{P}$ is given. Using the precedent chart let's admit that $a(P) \in H_M$ and in this case from the structure of the proposition P will be eliminated the hypotheses that contain the characteristic $\neg a$, so that the proposition P is equivalent to the proposition P_{AB}.

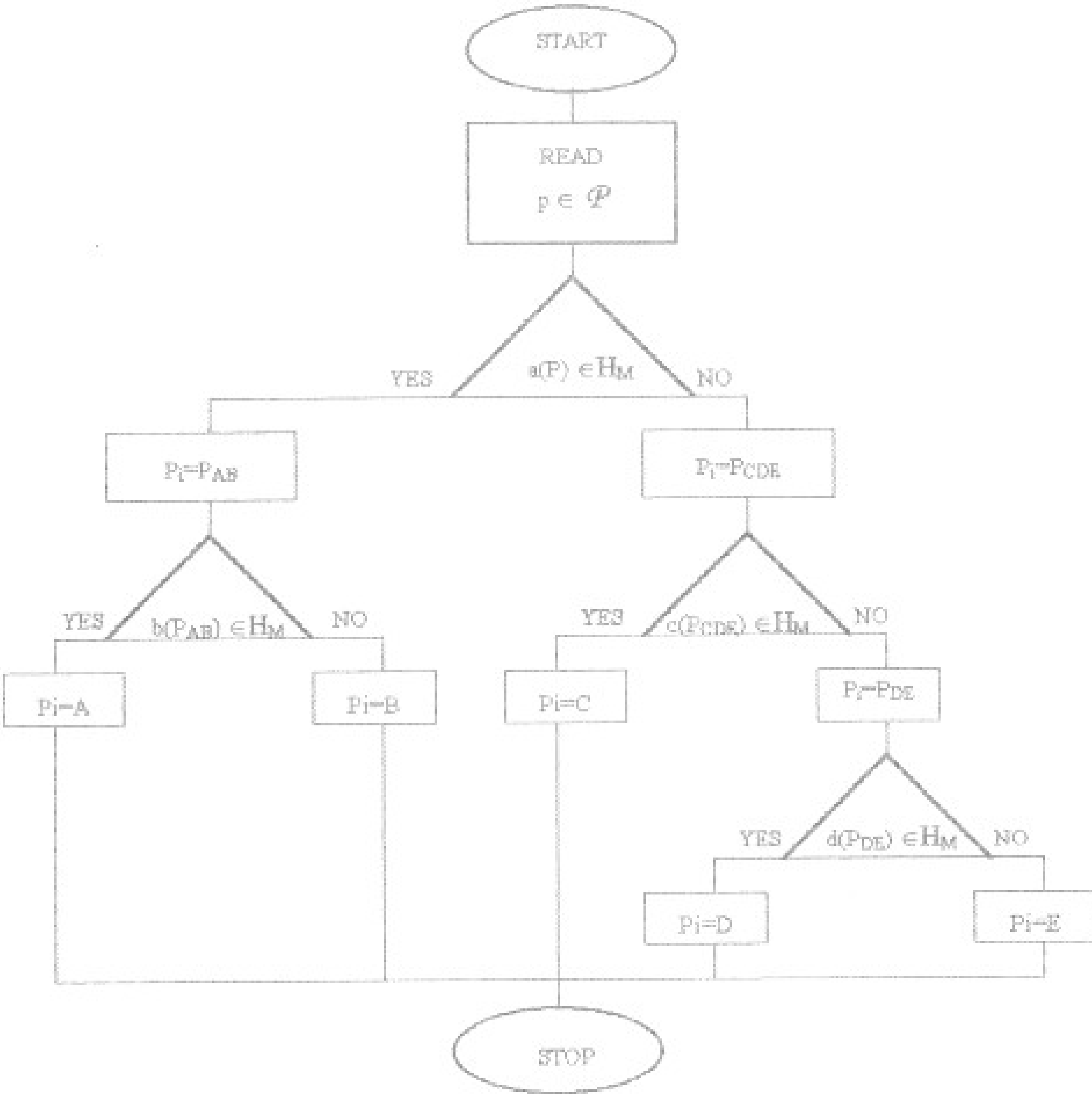
If $b(P_{AB}) \in H_M$ then from the structure of the proposition P_{AB} will be eliminated hypotheses that contain the characteristic $\neg b$ and so P_{AB} is equivalent to the proposition A, that is P is a proposition of the type A.

If $b(P_{AB}) \notin H_M$ from the structure of the proposition P_{AB} will be eliminated hypotheses that contain the characteristic b and so P_{AB} is equivalent to the proposition of the type B, that is P is a proposition of the type B.

If $a(P) \notin H_M$ then from the structure of the proposition P will be eliminated hypotheses that contain the characteristic a and consequently P is a proposition equivalent to the proposition P_{CDE} and if $c(P_{CDE}) \in H_M$ then from the structure of the proposition P_{CDE} will be eliminated hypotheses that contain the characteristic $\neg c$ and consequently the proposition P_{CDE} is equivalent to the proposition of the type C, that is P is equivalent to the proposition of the type C.

If $c(P_{CDE}) \notin H_M$ then out of the structure of the proposition P_{CDE} will be eliminated hypotheses that contain the characteristic c and consequently the proposition P_{CDE} is equivalent to the proposition P_{DE}. If $d(P_{DE}) \in H_M$ then from the structure of the proposition P_{DE} will be eliminated hypotheses that contain the characteristic $\neg d$ and consequently the proposition P_{DE} is equivalent to the proposition of the type D, and consequently the proposition P is equivalent to the proposition of the type D. If $d(P_{DE}) \notin H_M$ then from the proposition P_{DE} will be eliminated hypotheses that contain the characteristic d and consequently the proposition P_{DE} is equivalent to the proposition of the type E, and consequently the proposition P is equivalent to the proposition of the type E.

From the former analyses we obtain the following logical chart:



by which we can establish for a given proposition if it is of a type A, B, C, D or E.

If the process of learning it is necessary to go over to a higher level of study then the hypotheses A,B,C,D,E will change as well as the characteristics a,b,c,d but the method of analysis of the new level of study is the same with the one presented above.

In the process of learning an important problem is that to find the algorithm that would establish the less in time of the authenticity of the information taking into account that the multitude S of subjects liable to the education react in a different manner after the reception of the information when the system S has become the system S(t) that depends on the nature of information and on the way it is intercepted.

After a time t the system S(t) doesn't intercept the information that is transmitted to it and consequently it (the information) hasn't a real support, that is the message transmitted to the system S(t) can't constitute an information.

In this way a wear of information is produced that doesn't take place in the process of education but because of some actions that take place in the inner part (inward) of the system S(t). This phenomenon is produced because the information acts upon the system S after a time t has passed and S couldn't enlarge its capacity of perception of the new informations, consequently a loss of perception takes place and not a consumption of information.

The quality of transmission of an information of the system S is established by the error coefficient $k_e = N_e : N_T$ where N_e is the number of incorrect bits receipted and N_T is the number of transmitted bits.

Taking into account that the information is transmitted to the system S in bigger interval than the duration of a bit of information we deduce that k_e may be estimated through the probability P_e to recept an incorrect bit in the process of transmission. The probability to have error when an information is transmitted to a system S, is:

$$(32) \qquad P_{a,N} = \frac{1}{\sqrt{2\pi}} \int_z^{\infty} e^{-\frac{x^2}{2}} dx$$

where N is the level of the varying noise having the dispersion σ_N^2 and $z=b_0:N$, b_0 being the level of the symbol of information ([7]).

An important part in the establishing of the superior border, respectively of the inferior border of the probability $P_{a,N}$ is played by the function:

$$g_a : [0, \infty) \to [0, \infty),\ g_a(x) = e^{-ax^2} \int_0^x e^{t^2} dt$$

where $a \in R$, $a > 1$.

REFERENCE.

1. ALTAN, H., L'organisation biologique et la théorie d' information , Paris, 1972

2. CHERNOFF, H., The scoring of multiple choice questionnaires, Annals of athematics Statistics, 1962;

3. GOLDMAN, S., Information theory, London, 1953;

4. DIEUDONNÉ, J.

5. FISCHER, R.A., Probability, likelihood and quantity of information in the logic of uncertain inference. Proc. Royal A. 146, 1934, p.1-8;

6. HARTLEY, R.V.H., Transmission of Information. Bell System Technical Journal, 3, 1928, p.535-563;

7. LUKY, R.W., Principles of Data Communication. New York, Mc. Grau-Hill, 1962;

8. COURRY, HASKELL, Foundation of Mathematical Logic, New York, 1963;

9. ONICESCU, O., MIHOC, GH., Sur les thaines de variables statistiqe, Bull. Sci. Math. 59, 174-192(1935);

10. SHANNON, CL., A mathematical theory of communications. Bell. System Tech. J, 27 (1948);

Received 01.09.1996

Academia Trupelor de Uscat
str. Revoluţiei, 7
RO-2400 Sibiu
ROMANIA