Buletinul Ştiinţific al Universităţii "Politehnica" din Timişoara

Seria ELECTRONICĂ şi TELECOMUNICAŢII
TRANSACTIONS on ELECTRONICS and COMMUNICATIONS

Tom 51(65), Fascicola 2, 2006

On Semantic Feature of Information

Valeriu Munteanu[1], Daniela Tarniceriu[1]

Abstract – Beside the objective or quantitative characteristic of a message, appraised by the probability with which it is supplied, its semantic or qualitative characteristic, appraised by a certain utility or importance is, additionally, considered. In this paper we determine the quantitative – qualitative information, the quantitative – qualitative entropy of a discrete, complete and memoryless source as well as the main properties of the quantitative – qualitative entropy.
Keywords: semantic sources, entropies.

I. INTRODUCTION

In the elaboration of the information concept in classical sense, only its quantitative feature is considered, the semantic (qualitative) one being neglected [1, 2, 3]. In this case only the probabilities of random events are considered for computing the information. In cybernetic systems the transmitted information is used for a certain goal. Considering only the probabilistic dependencies between the transmitted messages is not enough, because the transmission efficiency also depends on the choosing of those messages that serve to the pursued goal. This is the reason why, in a cybernetic system, also the quality of the transmitted messages should be measured. A semantic source is characterized both quantitatively and qualitatively. Thus, the transmitted information will depend both on the objective part of the experiment through the event probabilities and on its utility or importance, that reflects the subjective part of the experiment related to a certain goal. Identical events, with same probabilities can have different utilities for different cybernetic systems, even if their goals are the same. On the other hand, for the same cybernetic system, the same events can have different utilities, when the goal is changed. Therefore, the utilities of different events are related both to the proposed goal and to the cybernetic system used to its achievement. In [4-8] authors have studied generalized coding theorems by considering different generalized measures. A first attempt to measure information both quantitatively and qualitatively is given in [9]. In [10] an introduction of the quantitative - qualitative information is presented. In this paper a fully presentation of semantic sources is performed, by deriving the quantitative – qualitative information, its average value for a discrete, complete and memoryless source, emphasizing the main properties of semantic sources.

II. DETERMINING THE QUANTITATIVE – QUALITATIVE INFORMATION

Let S be a discrete, complete and memoryless source characterized by the distribution:

$$S:\begin{pmatrix} s_1 & s_2 & \cdots & s_n \\ p_1 & p_2 & \cdots & p_n \\ u_1 & u_2 & \cdots & u_n \end{pmatrix} \qquad (1)$$

where

- s_i represents the source messages, that is, signals corresponding to ideas, images, data which has to be transmitted to a correspondent;
- p_i denotes the probabilities with which the source delivers its messages, so that

$$0 \le p_i \le 1,\ i = 1,\dots,n\ , \qquad (2)$$

$$\sum_{i=1}^{n} p_i = 1 \qquad (3)$$

- u_i denotes the utility or importance of the message s_i. This utility is appraised by a positive real number which reflects the semantic characteristics of the message as function of the given goal and of the system.

Theorem 1

The quantitative – qualitative information or the semantic information attached to the message s_k, denoted by $i_{pu}(s_k)$, is computed by

$$i_{pu}(s_k) = b\log_\alpha p_k + au_k\ , \qquad (4)$$

where $\alpha \in \{0,1\} \cup \{1,\infty\}$; $a,b,u_i \in R$.

Proof

The quantitative – qualitative information attached to the message s_i will be a function *F*, of the probability p_i and the utility u_i, respectively. To derive this function, two messages, independent both probabilistic and logic – causal, s_i and s_j, are

[1] Facultatea de Electronică şi Telecomunicaţii, Departamentul Comunicaţii Bd. Carol I, nr.11, 700506 Iasi, e-mail vmuntean@etc.tuiasi.ro

considered. Their probabilities are p_i and p_j, respectively and the utilities u_i and u_j, respectively.

If

$$\sigma_k = s_i s_j \ , \tag{5}$$

$$p(\sigma_k) = p_i p_j \tag{6}$$

$$u(\sigma_k) = u_i + u_j \tag{7}$$

The quantitative – qualitative information the message σ_k can deliver is, generally, a function of the probability $p(\sigma_k)$ and the utility $u(\sigma_k)$, as

$$i_{pu}(\sigma_k) = F[p(\sigma_k), u(\sigma_k)] \tag{8}$$

where F is a function to be found out.

We also assume that the quantitative – qualitative information given by two events independent both statistical and causal is equal to the sum of two pieces of information attached to each of them. Therefore,

$$i_{pu}(\sigma_k) = i_{pu}(s_i) + i_{pu}(s_j) \tag{9}$$

Considering (8), (9) becomes

$$F[p(\sigma_k), u(\sigma_k)] = F(p_i, u_i) + F(p_j, u_j) \tag{10}$$

with

$$0 \le p_i \le 1; 0 \le p_j \le 1; u_i, u_j \in R \tag{11}$$

In order to determine the solution of the functional equation (10), the following functions are defined

$$z_i = \log_\alpha p_i, z_i \in R, \tag{12}$$

$$z_j = \log_\alpha p_j, z_j \in R \tag{13}$$

where $\alpha \in \{0,1\} \cup \{1,\infty\}$.

Using (12) and (13), (10) becomes

$$F(\alpha^{z_i+z_j}, u_i + u_j) = F(\alpha^{z_i}, u_i) + F(\alpha^{z_j}, u_j) \tag{14}$$

Denoting

$$G(z,u) = F(\alpha^z, u) \tag{15}$$

equation (14) can be written as

$$G(z_i + z_j, u_i + u_j) = G(z_i, u_i) + G(z_j, u_j) \tag{16}$$

For $z_i = z_j = 0$, we have

$$G(0, u_i + u_j) = G(0, u_i) + G(0, u_j) \tag{17}$$

Denoting

$$f(u) = G(0, u) \tag{18}$$

equation (17) becomes

$$f(u_i + u_j) = f(u_i) + f(u_j). \tag{19}$$

The solution of this functional equation is given in [11], provided the function $f(u)$ is continuous at least in one point, being of the form

$$f(u) = a \cdot u, (\forall) u \in R, a \in R. \tag{20}$$

For $u_i = u_j = 0$, from (16) we obtain

$$G(z_i + z_j, 0) = G(z_i, 0) + G(z_j, 0) \tag{21}$$

Denoting by

$$g(z) = G(z, 0) \tag{22}$$

equation (17) becomes

$$g(z_i + z_j) = f(z_i) + f(z_j). \tag{23}$$

Its solution is

$$g(z) = b \cdot z, (\forall) z \in R, b \in R. \tag{24}$$

The functional equation (24) admits a solution, if $g(z)$ is continuous at least in one point.

For $u_i = z_j = 0$, from (16) we have

$$G(z_i, u_j) = G(z_i, 0) + G(0, u_j) \tag{25}$$

Making use of (18) and (22), we can write

$$G(z,u) = g(z) + f(u) \tag{26}$$

With (20) and (24), (26) becomes

$$G(z,u) = b \cdot z + a \cdot u, \quad (\forall) z \in R, (\forall) u \in R, a \in R, b \in R \tag{27}$$

or, considering (12), (13) and (15), we get the solution of the functional equation (21), as

$$F(p,u) = b \cdot \log_\alpha p + a \cdot u \tag{28}$$

According to (8) and (28), the quantitative - qualitative information attached to the message s_k, with the utility u_k and delivered with the probability p_k, can be computed by (4).

The neglect of the qualitative characteristic of information consists in removing the second term in the right hand of (4). In this way, we get the classical calculus relation for the quantitative information attached to a message [1, 2, 3].

III. DETERMINING THE AVERAGE QUANTITATIVE – QUALITATIVE INFORMATION FOR A SEMANTIC MEMORYLESS SOURCE

Let S be a discrete, complete and memoryless source characterized by the distribution given in (1).

Theorem 2

The average quantitative – qualitative information is computed by

$$H_{pu}(S) = b\sum_{k=1}^{n} p_k \log_\alpha p_k + a\sum_{k=1}^{n} p_k u_k \tag{29}$$

Proof

The quantitative – qualitative information $i_{pu}(s_k)$ defined in (4) determines a discrete random variable, which takes on values with probabilities $p_k, k = 1,2,...,n$. The average value of this information, called quantitative – qualitative entropy or semantic entropy of the source S and denoted by $H_{pu}(S)$, can be computed by

$$H_{pu}(S) = \sum_{k=1}^{n} p_k i_{pu}(s_k) \tag{30}$$

or, considering (4), we obtain (29).

If $u_1 = u_2 = ... = u_n = 0$, that is, the qualitative characteristic is neglected, (29) becomes

$$H_{pu}(S) = b\sum_{k=1}^{n} p_k \log_\alpha p_k \tag{31}$$

In order to obtain the entropy defined by Shannon [2], we set

$$b = -1, \alpha = 2. \tag{32}$$

With (32), (29) becomes

$$H_{pu}(S) = -\sum_{k=1}^{n} p_k \log_2 p_k + a\sum_{k=1}^{n} p_k u_k . \quad (33)$$

If the total utility is represented by *U*, that is

$$\sum_{k=1}^{n} u_k = U , \quad (34)$$

then, for $p_1 = p_2 = ... = p_n = \frac{1}{n}$, we can write

$$a\sum_{k=1}^{n} p_k u_k = \frac{aU}{n} . \quad (35)$$

But $\frac{aU}{n}$ measures the average utility per message for equally likely probabilities, which is equal to $\frac{U}{n}$. It follows that *a* has to be equal to unity.

In the sequel the quantitative – qualitative entropy or the semantic entropy will be computed by

$$H_{pu}(S) = -\sum_{k=1}^{n} p_k \log_2 p_k + \sum_{k=1}^{n} p_k u_k \quad (36)$$

and the quantitative - qualitative information or semantic information by

$$i_{pu}(s_k) = -\log_2 p_k + u_k . \quad (37)$$

IV. MAIN PROPERTIES OF THE QUANTITATIVE – QUALITATIVE ENTROPY

Property 1: If $u_k \in R^+$, the quantitative – qualitative entropy is nonnegative, i.e.

$$H_{pu}(S) \geq 0 \quad (38)$$

Property 2: If $p_k = 1, u_k = 0;\ 1 \leq k \leq n$, then

$$H_{pu}(S) = 0 \quad (39)$$

Property 3: The maximum value of the entropy with respect to the probabilities p_k, for given utilities, is:

$$\max_{p_k} H_{pu}(S) = \log_2\left(\sum_{k=1}^{n} 2^{u_k}\right) \quad (40)$$

This maximum entropy will be denoted by $H_m(S)$.

Proof

Let

$$\phi(p_1, p_2, ..., p_n; u_1, u_2, ..., u_n) = -\sum_{k=1}^{n} p_k \log_2 p_k + \sum_{k=1}^{n} p_k u_k + \lambda(\sum_{k=1}^{n} p_k - 1) \quad (41)$$

where λ is a real positive number (the Lagrange multiplier).

The extreme of the function ϕ with respect to p_k for u_k fixed, coincides with the extreme of the function $H_{pu}(S)$. The necessary condition of extreme is given by the system

$$\frac{\partial\phi(p_1, p_2, ..., p_n; u_1, u_2, ..., u_n)}{\partial p_k} = 0, 1 \leq k \leq n \quad (42)$$

or, equivalently, taking into account (41),

$$-\log_2 p_k - \log_2 e + u_k + \lambda = 0, 1 \leq k \leq n , \quad (43)$$

from where

$$p_k = 2^{\lambda + u_k - \log_2 e} \quad (44)$$

Since

$$\sum_{k=1}^{n} p_k = 1 \Rightarrow \lambda = \log_2 e - \log_2(\sum_{k=1}^{n} 2^{u_k}) . \quad (45)$$

Substituting (45) into (44), we have

$$p_k = \frac{2^{u_k}}{\sum_{k=1}^{n} 2^{u_k}}, k = 1, 2, ..., n . \quad (46)$$

Finally, substituting (46) into (36), relation (40) follows. It is easy to prove that this extreme is a maximum one.

Property 4

The absolute maximum value of entropy, denoted by $H_{ma}(S)$ is given by

$$H_{ma}(S) = \log_2 n + \frac{U}{n} \quad (47)$$

Proof

We want to find the utilities the messages s_k have to possess, so that

$$\left(\sum_{k=1}^{n} 2^{u_k}\right) = \max , \quad (48)$$

under the constraint of (34).

If (48) is satisfied, then the entropy computed by (40) becomes an absolute maximum one.

In order to determine the utilities u_k for which (48) is satisfied, the Lagrange multipliers method is used. The following function will be constructed:

$$\Psi(u_k) = \sum_{k=1}^{n} 2^{u_k} + \lambda\left(U - \sum_{k=1}^{n} u_k\right), \lambda > 0 , \quad (49)$$

where λ is the Lagrange multiplier. The maximum value of $\Psi(u_k)$ is attained simultaneously with the maximum value of (48).

The necessary condition for extreme is

$$\frac{\partial\Psi(u_k)}{\partial u_k} = 0, k = 1, ..., n \quad (50)$$

Substituting (49) into (50), we have

$$u_k = \log_2\left(\frac{\lambda}{\ln 2}\right), k = 1, 2, ..., n \quad (51)$$

Since

$$\sum_{k=1}^{n} u_k = U \quad (52)$$

it results

$$\lambda = 2^{U/n} \ln 2 \quad (53)$$

Substituting (53) into (51), we get

$$u_k = \frac{U}{n}, k = 1, 2, ..., n \quad (54)$$

We can easily prove that the so obtained extreme value is a maximum one, because

$$\frac{\partial^2\Psi(u_k)}{\partial u_k^2} > 0, \frac{\partial^2\Psi(u_k)}{\partial u_k \partial u_i} = 0,\ i, k = 1, ..., n, i \neq k . \quad (55)$$

According to (54), the quantitative – qualitative entropy is maximized when all source messages have identical qualitative weights (the same utilities), equal to the arithmetic mean value of the utilities.

Substituting (54) into (46), we get

$$p_k = \frac{1}{n}, k = 1, ..., n . \quad (56)$$

Substituting (54) and (56) into (36), (47) results.

V. CONCLUSIONS

In this paper, besides the objective or quantitative characteristic of a message, appraised by the probability with which it is supplied, its semantic or qualitative characteristic, appraised by a certain utility is, additionally, considered. The quantitative – qualitative information attached to a message (eq. 4), as well as the quantitative – qualitative entropies of a discrete, complete and memoryless source (eq. 29) are derived. The maximum value of the entropy of a semantic source is determined with respect to the probabilities the messages are delivered. The constraints on the utilities and probabilities for obtaining the absolute maximum entropy of a semantic source are inferred. These relations represent generalizations of the classical concepts on information. The quantitative – qualitative information results as the sum between a quantitative information and a qualitative one. If the qualitative characteristic is neglected, by dropping out the second term in relations (4) and (29), the classical known relations [2] are obtained. When only the qualitative characteristic is required, the first term in the relations above is dropped out. The main properties of entropy for semantic sources are also established. The semantic entropy defined in this paper preserves the properties of the classical entropy defined by Shannon.

REFERENCES

[1] G. R.Gallager, *Information theory and reliable communications*, New York, John Wiley and Sons Inc., 1968.

[2] C. E.Shannon, "A mathematical theory of communication", BSTJ, 27 1948, pp. 379-423, 623-656.

[3] T.Cover, J. A.Thomas, *Elements of Information Theory*, Wiley, 1991.

[4] G. Longo, "A noisless coding theorem for sources having utilities", *SIAM J. Appl. Math*., 30 (4), 1976, pp. 739-748.

[5] A.Gurdial, F. Pessoa, "On useful information of order "α . *J Comb. Information and Syst. Sci*., 2, 1977, pp. 158-162.

[6] A. B. Khan, R. Autar, "On useful information of order α and β ", *Soochow J. Math.*, 5, 1979, pp. 93-99.

[7] R. Autar, A. B. Khan, "On generalized useful information for incomplete distribution", *J. of Comb. Information and Syst. Sci*., 14 (4), 1989, pp. 187-191.

[8] A. B. Khan, B. A.Bhat, S. Pirzada, "Some results on a generalized useful information measure", *J. Inequal. Pure and Appl. Math*, 6 (4), 2005, pp. 1-5.

[9] M.Belis, S. Guiasu, "A quantitative-qualitative measure of information in cybernetics", *IEEE Trans. Inf. Theory* IT – 14, 1968, pp. 593-594.

[10] V. Munteanu, P. Cotae, "The entropy with preference", *Int. J. Electronics and Comm. AEU*, 46, 1992, pp. 429 – 431.

[11] G. M. Fihtengol'c, *Differential and integral calculus*, Ed. Tehnica, Bucharest, 1964.