Mathematical Problems of Computer Science 21, 2000, 77-82.

Reliability Criterion in Successive Refinement

Ashot N. Harutyunyan*

Institute for Informatics and Automation Problems of NAS RA and YSU
E-mail: ashar@ipia.sci.am

Abstract

We study the *reliability* criterion in the concept of source divisibility named also successive refinement of information. Messages of a source must be coded for transmission to receiver with distortion not exceeding Δ_1 with error probability exponent (reliability) E_1, and then, using an auxiliary information, restated with a more precise distortion $\Delta_2 \leq \Delta_1$ with error probability exponent $E_2 \geq E_1$. Let $R(E, \Delta)$ be rate-reliability-distortion function and R' be the rate of the additional encoding. Successive refinement of information (or divisibility of source) with reliability requirement from (E_1, Δ_1) to (E_2, Δ_2) is possible provided that $R(E_2, \Delta_2) = R(E_1, \Delta_1) + R'$.

We present a condition necessary and sufficient for the successive refinement with given reliability $E_1 = E_2 = E$.

1 Introduction

The concept of source *divisibility* was introduced by Koshelev [1]-[3] as a criterion of efficiency for source coding multilevel system. The same concept was independently redefined by Equitz and Cover in [5] and named *"successive refinement of information"*. The idea is to achieve the rate-distortion limit at each level of successively more precise transmission. We consider a more exact formulation of the problem in the simple case of two levels.

Assume that we transmit information to two users, the requirement of the first on distortion is no larger than Δ_1 and demand of the second user is more accurate: $\Delta_2 \leq \Delta_1$. It is well known from rate-distortion theory that the value $R(\Delta_1)$ of the rate-distortion function is the minimal satisfactory transmission rate at the first destination. Adding an information of a rate R' addressed to the second user the fidelity can be made more precise providing no larger distortion than Δ_2. It is interesting to know when it is possible to do so that $R(\Delta_1) + R' = R(\Delta_2)$. The answer to this question is given in the works of Koshelev [1]-[4], and of Equitz and Cover [5]. Koshelev argued that Markov condtion for random variables (RV) characterizing the system is sufficient to achieve rate-distortion limit and later sufficiency and also necessity of the condition was established in the paper of Equitz and Cover [5]. An other proof of the result and an interpretation of Markov condition are given in the work of Rimoldi [6].

*The work of the author was supported in part by INTAS under Young Scientists Fellowship Grant 00-4162.

By analogy with the characterization of successive refinement with the help of the rate-distortion function we operate with terms of the rate-reliability-distortion function (see [7]) and prove corresponding general condition neccessary and sufficient for possibility of successive refinement with reliability requirement.

The description of the source coding system and a preliminary result are given in sections 2 and 3, in section 4 the general definition of the successive refinement and in section 5 our result (Theorem 3) and its proof are presented.

2 The communication system

Let the probability distribution (PD) of messages of the discrete memoryless source (DMS) $\{X\}$ of finite alphabet $\mathcal{X}$ is $P^* = \{P^*(x), x \in \mathcal{X}\}$. Reproduction alphabets of two receivers accordingly are $\mathcal{X}^1$ and $\mathcal{X}^2$ and the corresponding single-letter distortion measures are

$$d_k : \mathcal{X} \times \mathcal{X}^k \to [0; \infty),\ k = 1, 2.$$

Distortions $d_k(\mathbf{x}, \mathbf{x}^k)$ $(k = 1, 2)$ between a source N-length message $\mathbf{x}$ and its reproducted versions $\mathbf{x}^k$ are considered as averages of per-letter distortions.

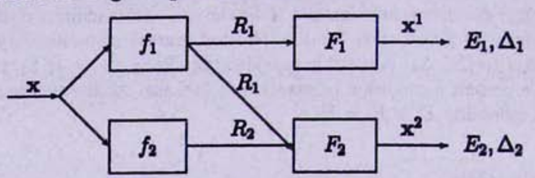

Fig. 1. Two-level communication system.

A code $(f, F) = (f_1, f_2, F_1, F_2)$ for the system consists of two encoders:

$$f_k : \mathcal{X}^N \to \{1, 2, ..., L_k(N)\},\ k = 1, 2,$$

and two decoders:

$$F_1 : \{1, 2, ..., L_1(N)\} \to (\mathcal{X}^1)^N,$$
$$F_2 : \{1, 2, ..., L_1(N)\} \times \{1, 2, ..., L_2(N)\} \to (\mathcal{X}^2)^N.$$

Let's consider the following sets:

$$\mathcal{A}_1 = \{\mathbf{x} \in \mathcal{X}^N : F_1(f_1(\mathbf{x})) = \mathbf{x}^1, d^1(\mathbf{x}, \mathbf{x}^1) \le \Delta_1\},$$
$$\mathcal{A}_2 = \{\mathbf{x} \in \mathcal{X}^N : F_2(f_1(\mathbf{x}), f_2(\mathbf{x})) = \mathbf{x}^2, d^2(\mathbf{x}, \mathbf{x}^2) \le \Delta_2\}.$$

The probability of error (for a code (f, F)) at each destination is

$$e_k(f, F, \Delta_k, N) = 1 - P^*(\mathcal{A}_k),\ k = 1, 2.$$

Let $\mathbf{E} = (E_1, E_2)$, $\Delta = (\Delta_1, \Delta_2)$. We say that (R_1, R_2) $(R_k \ge 0, k = 1, 2)$ is a $(\mathbf{E}, \Delta)$-achievable pair of rates for $E_k > 0$, $\Delta_k \ge 0$, $k = 1, 2$, if for every $\epsilon > 0$ and sufficiently large N there exists a code (f, F), such that

$$\frac{1}{N} \log L_k(N) \le R_k + \epsilon,$$

$$e_k(f, F, \Delta_k, N) \le \exp\{-NE_k\}, k = 1, 2.$$

E_k ($k = 1, 2$) are called *reliabilities*. Let's denote the set of $(\mathbf{E}, \boldsymbol{\Delta})$-achievable rates for the system by $\mathcal{R}(\mathbf{E}, \boldsymbol{\Delta})$.

Let $P = \{P(x), x \in \mathcal{X}\}$ be a PD on $\mathcal{X}$ and

$$Q = \{Q(x^1, x^2|x),\ x \in \mathcal{X},\ x^k \in \mathcal{X}^k,\ k = 1, 2\}$$

be a conditional PD on $\mathcal{X}^1 \times \mathcal{X}^2$. Denote by $D(P \parallel P^*)$ the divergence between PD P and P^*:

$$D(P \parallel P^*) = \sum_x P(x) \log \frac{P(x)}{P^*(x)}.$$

And let

$$\alpha(E_k) = \{P : D(P \parallel P^*) \le E_k\},\ k = 1, 2.$$

In this article we shall consider the case of equal reliabilities, i.e. $E_1 = E_2 = E$.

Consider a function $\Phi(P, \mathbf{E}, \boldsymbol{\Delta})$, values of which are conditional PD Q corresponding to a PD P such that for a given $\boldsymbol{\Delta}$ if $P \in \alpha(E)$ then $E_{P,Q} d_k(X, X^k) \le \Delta_k$, $k = 1, 2$.

Let $\mathcal{M}(P, \mathbf{E}, \boldsymbol{\Delta})$ be the collection of all such functions $\Phi(P, \mathbf{E}, \boldsymbol{\Delta})$ for given $\mathbf{E}, \boldsymbol{\Delta}$ and P.

3 Achievable rates region

In the paper of Maroutian [8] the region $\mathcal{R}(\mathbf{E}, \boldsymbol{\Delta})$ in the case of $E_1 \ge E_2$ is found. In Theorem 1 we describe it for the case $E_1 = E_2 = E$ for the formulation of our result (Theorem 3).

Let

$$I_{P,Q}(X \wedge X^1, X^2) = \sum_{x, x^1, x^2} P(x) Q(x^1, x^2 \mid x) \log \frac{Q(x^1, x^2 \mid x)}{\sum_x P(x) Q(x^1, x^2 \mid x)}$$

be the Shannon mutual information between RVs X and X^1, X^2 defined by PDs P and Q.

The separate informations $I_{P,Q}(X \wedge X^k)$, $k = 1, 2$, between the RVs X and X^k are defined similarly using the marginal distributions $Q(x^k \mid x)$:

$$\sum_{x^j, j \ne k} Q(x^j, x^k \mid x) = Q(x^k \mid x),\ j, k = 1, 2.$$

Theorem 1. For every $0 < E_1 = E_2 = E$ and $0 \le \Delta_2 \le \Delta_1$

$$\mathcal{R}(\mathbf{E}, \boldsymbol{\Delta}) = \bigcap_{P \in \alpha(E)} \bigcup_{\Phi(P, \mathbf{E}, \boldsymbol{\Delta}) \in \mathcal{M}(P, \mathbf{E}, \boldsymbol{\Delta})} \{(R_1, R_2) : R_1 \ge I_{P, \Phi(P, \mathbf{E}, \boldsymbol{\Delta})}(X \wedge X^1),$$

$$R_1 + R_2 \ge I_{P, \Phi(P, \mathbf{E}, \boldsymbol{\Delta})}(X \wedge X^1 X^2)\}.$$

4 Successive refinement with reliability

Denote by $R(E_k, \Delta_k)$ the rate-reliability-distortion functions for each level of requirements of users on reliability E_k and distortion Δ_k. And let $R(\Delta_k)$ be the corresponding rate-distortion functions ($k = 1, 2$). It is known [7] that

$$R(E_k, \Delta_k) = \max_{P : P \in \alpha(E_k)} \min_{Q : E_{P,Q} d_k(X, X^k) \le \Delta_k} I_{P,Q}(X \wedge X^k), \tag{1}$$

and [9]

$$R(\Delta_k) = \min_{Q: E_{P,Q} d_k(X,X^k) \leq \Delta_k} I_{P,Q}(X \wedge X^k), \ k = 1,2 \tag{2}$$

Definition. We say that successive refinement from (E_1, Δ_1) to (E_2, Δ_2), for $\Delta_1 \geq \Delta_2$ and $E_1 \leq E_2$ (or source divisibility) is possible if the pair of rates

$$(R(E_1, \Delta_1), R(E_2, \Delta_2) - R(E_1, \Delta_1))$$

is $(\mathbf{E}, \Delta)$-achievable for the considered multilevel system.

In our notations the result of Equitz and Cover may be formulated as follows:

Theorem 2 (Equitz and Cover [5]). For the DMS with distribution P^* and distortion measure $d_1 = d_2 = d$, the pair $(R(\Delta_1), R(\Delta_2) - R(\Delta_1))$ is achievable iff there exists a conditional probability Q, such that

$$R(\Delta_1) = I_{P^*,Q}(X \wedge X^1), \ E_{P^*,Q} d(X, X^1) \leq \Delta_1,$$

$$R(\Delta_2) = I_{P^*,Q}(X \wedge X^2), \ E_{P^*,Q} d(X, X^2) \leq \Delta_2,$$

and X, X^2, X^1 form a Markov chain in that order.

5 The condition and proof

Theorem 3. For the DMS $\{X\}$ with generic distribution P^* and distortion measures $d_1 = d_2 = d$ the pair $(R(E, \Delta_1), R(E, \Delta_2) - R(E, \Delta_1))$ is $(\mathbf{E}, \Delta)$-achievable iff there exist pairs of PDs $P_1 \in \alpha(E), Q_1 \in \mathcal{M}(P_1, \mathbf{E}, \Delta)$ and $P_2 \in \alpha(E), Q_2 \in \mathcal{M}(P_2, \mathbf{E}, \Delta)$, such that

$$R(E, \Delta_1) = I_{P_1,Q_1}(X \wedge X^1),$$

$$R(E, \Delta_2) = I_{P_2,Q_2}(X \wedge X^2),$$

and RV X, X^2, X^1 form a Markov chain $X_{P_2} \rightarrow X^2 \rightarrow X^1$, where X_{P_2} is the RV X with the distribution P_2.

Corollary. In the case when the users requirements on the reliability are not present the result of Equitz and Cover (Theorem 2) formulated above follows from Theorem 3 by $E \rightarrow 0$.

Proof of Theorem 3.

Necessity. Assume that the pair

$$(R(E, \Delta_1), R(E, \Delta_2) - R(E, \Delta_1))$$

is $(\mathbf{E}, \Delta)$-achievable for the system. It follows from Theorem 1 that for each $P \in \alpha(E)$ and every $Q \in \mathcal{M}(P, \mathbf{E}, \Delta)$ next two inequalities are valid:

$$R(E, \Delta_1) \geq I_{P,Q}(X \wedge X^1), \tag{3}$$

$$R(E, \Delta_1) + R(E, \Delta_2) - R(E, \Delta_1) \geq I_{P,Q}(X \wedge X^1 X^2). \tag{4}$$

Then it follows from (3) and (1) that there exists a pair of distributions $P_1 \in \alpha(E)$ and corresponding $Q_1 \in \mathcal{M}(P_1, E, \Delta)$ such that

$$R(E, \Delta_1) = I_{P_1,Q_1}(X \wedge X^1).$$

From (4) we have for each $P \in \alpha(E)$ and for every $Q \in \mathcal{M}(P, E, \Delta)$:

$$R(E, \Delta_2) \geq I_{P,Q}(X \wedge X^1X^2) \geq I_{P,Q}(X \wedge X^2),$$

then among these distribitions there exist such P_2, Q_2 that

$$R(E, \Delta_2) \geq I_{P_2,Q_2}(X \wedge X^1X^2) \geq I_{P_2,Q_2}(X \wedge X^2) = R(E, \Delta_2).$$

This yields

$$I_{P_2,Q_2}(X \wedge X^1X^2) = I_{P_2,Q_2}(X \wedge X^2),$$

which is equivalent to $X_{P_2} \rightarrow X^2 \rightarrow X^1$ the Markovian condition for RVs X, X^2 and X^1. The proof of the necessity is complete.

Sufficiency. Let $P_1 \in \alpha(E)$, $Q_1 \in \mathcal{M}(P_1, \mathbf{E}, \mathbf{\Delta})$ and $P_2 \in \alpha(E)$, $Q_2 \in \mathcal{M}(P_2, \mathbf{E}, \mathbf{\Delta})$ be such PDs that

$$R(E, \Delta_1) = I_{P_1,Q_1}(X \wedge X^1), \tag{5}$$

$$R(E, \Delta_2) = I_{P_2,Q_2}(X \wedge X^2), \tag{6}$$

and

$$X_{P_2} \rightarrow X^2 \rightarrow X^1. \tag{7}$$

We have to prove that then the pair of rates $(R(E, \Delta_1), R(E, \Delta_2) - R(E, \Delta_1))$ is $(\mathbf{E}, \mathbf{\Delta})$-achievable.

Taking into account (1), the Markovian condition (7), the equality (6) can be rewritten as follows

$$R(E, \Delta_2) = I_{P_2,Q_2}(X \wedge X^2) = I_{P_2,Q_2}(X \wedge X^1X^2) =$$

$$= \max_{P: P \in \alpha(E)} \min_{Q: E_{P,Q} d_2(X, X^2) \leq \Delta_2,\ X_P \rightarrow X^2 \rightarrow X^1} I_{P,Q}(X \wedge X^1X^2) \geq$$

$$\geq \max_{P: P \in \alpha(E)} \min_{Q: E_{P,Q} d_2(X, X^2) \leq \Delta_2,} I_{P,Q}(X \wedge X^1X^2).$$

The assumption (5) and the last inequality allow to conclude that

$$(R(E, \Delta_1), R(E, \Delta_2) - R(E, \Delta_1)) \in \mathcal{R}(\mathbf{E}, \mathbf{\Delta}).$$

The proof of the sufficiency and the theorem is complete.

Remarks. We are sure that Theorem 3 can be strengthen for the case of different reliabilities, i.e. $E_1 \neq E_2$, by proving a relevant (when $E_1 < E_2$) result for rates-reliabilities-distortions region (Theorem 2), in other words by generalizing the result of Maroutian [8].

In the work of Equitz and Cover [5] various examples of successively refinable sources are presented. It is natural to establish the possibility of successive refinement under reliability criterion also by construction of an example.

We aimed to clarify these questions in the future investigations.

References

[1] V. N. Koshelev, "Multilevel source coding and data-transmission theorem", in *Proc. VII All-Union Conf. on Theory of Coding and Data Transm.*, Vilnius, U.S.S.R., pt. 1, pp. 85–92, 1978.

[2] V. N. Koshelev, "Hierarchical coding of discrete sources", *Problemy peredachi informatsii*, vol. 16, no. 3, pp. 31–49, (in Russian), 1980.

[3] V. N. Koshelev, "An evaluation of the average distortion for discrete scheme of sequential approximation", *Problemy peredachi informatsii*, vol. 17, no. 3, pp. 20–30, 1981.

[4] V. N. Koshelev, "On divisibility of discrete sources with the single-letter-additive measure of distortion", *Problemy peredachi informatsii*, vol. 30, no. 1, pp. 31–50, (in Russian), 1994.

[5] W. H. R. Equitz and T. M. Cover, "Successive refinement of information", *IEEE Trans. on Inform Theory*, vol. IT–37, no. 2, pp. 269–275, March 1991.

[6] B. Rimoldi, "Successive refinement of information: Characterization of the achievable rates", *IEEE Trans. on Inform Theory*, vol. IT–40, no. 1, pp. 253–259, January 1991.

[7] E. A. Haroutunian and B. Mekoush, "Estimates of optimal rates of codes with given error probability exponent for certain sources", (in Russian), Abstr. of Sixth Int. Symposium on Inform. Theory, Tashkent, vol. 1, pp. 22–23, 1984.

[8] R. Sh. Maroutian, "Achievable rates for multiple descriptions with given exponent and distortion levels", *Problemy peredachi informatsii*, vol. 26, no. 1, pp. 83–89, (in Russian), 1990.

[9] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems*, New York: Academic, 1981.