ADVANCES IN APPLIED MATHEMATICS 10, 75–94 (1989)

A General 4-Words Inequality with Consequences for 2-Way Communication Complexity

R. AHLSWEDE, N. CAI, AND Z. ZHANG

Universität Bielefeld, Fakultät für Mathematik, Postfach 8640, 4800 Bielefeld, Federal Republic of Germany

1. INTRODUCTION

The inequality of [1] (see also [2, 3]) was discovered in the analysis of the two-way complexity [5] of the Hamming distance function. Further analysis has led to much more general inequalities, culminating in the "4-words inequality" of [4], where also a program for further studies has been formulated. The papers [7, 8] make contributions to this program. Whereas in [8] earlier results of [1] are sharpened to the case of constant (resp. constant parity) Hamming distances, in [8] a 4-words inequality for another metric, namely the Lee metric, is proved.

The first result of the present paper is a very general 4-words inequality for arbitrary additive distortion functions (instead of distances). The 4-words property used seems to be so natural that we believe our new 4-words inequality (Theorem 1) to be in final form. Several special cases are stated as corollaries.

We also settle the constant union problem in an asymptotic sense via an exact solution of a new quantitative 1-sided constant union problem (Theorems 2, 3).

Another seemingly basic observation is the Decomposition Lemma in Section 7, which enables us to extend the asymptotic solution of the constant union problem to any sum-type function (Theorem 4). For these functions we give an upper bound for the 2-way communication complexity (Theorem 5) and we formulate a rule saying when this bound is tight. From Theorem 1 we get a lower bound, which often can be evaluated.

Many results still hold, if cardinalities are replaced by probabilities. This has been shown for the 4-words inequality (Theorem 1′) and a 4-word type generalization of the 1-sided constant union problem (Theorem 2′). The

0196-8858/89 $7.50

probability distributions are all assumed to be of product type. There is much work left to do for more general distributions.

2. A General 4-Words Inequality

Let $\mathcal{X}$ and $\mathcal{Y}$ be two finite sets. We consider functions

$$f: \mathcal{X} \times \mathcal{Y} \to \mathbb{Z}, \tag{2.1}$$

where $\mathbb{Z}$ is the ring of integers. With f we associate the "sum-type" function f_n: $\mathcal{X}^n \times \mathcal{Y}^n \to \mathbb{Z}$ defined by

$$f_n(x^n, y^n) = \sum_{t=1}^{n} f(x_t, y_t) \tag{2.2}$$

for all $x^n = (x_1, \ldots, x_n) \in \mathcal{X}^n$ and $y^n = (y_1, \ldots, y_n) \in \mathcal{Y}^n$.

We say that the pair (A, B) with $A \subset \mathcal{X}^n$ and $B \subset \mathcal{Y}^n$ satisfies the $\mathcal{R}$-4-words property, if

$$f_n(a^n, b^n) - f_n(a^n, b'^n) + f_n(a'^n, b'^n) - f_n(a'^n, b^n) \in \mathcal{R}. \tag{2.3}$$

Let $\mathcal{P}(f, \mathcal{R}, n)$ be the set of those pairs. We are interested in

$$M(f, \mathcal{R}, n) = \max\{|A||B|: (A, B) \in \mathcal{P}(f, \mathcal{R}, n)\}. \tag{2.4}$$

Let $\mathcal{P}^*(f, \mathcal{R}, n)$ be the set of those pairs in $\mathcal{P}(f, \mathcal{R}, n)$ assuming the maximal value $M(f, \mathcal{R}, n)$. Our first basic result can now be stated.

Theorem 1 (General 4-words inequality). *For any* $\mathcal{R} \subset \mathbb{Z}$

$$M(f, \mathcal{R}, n) \le M(f, \mathcal{R}, 1)^n. \tag{i}$$

Furthermore, if $0 \in \mathcal{R}$ *and* $M(f, \{0\}, 1) = M(f, \mathcal{R}, 1)$, *then equality holds in* (i).

Our proof below proceeds by induction on n and is based on two lemmas, which we first state and prove. We need a few definitions.

If C is a set of sequences of length n over some alphabet, then we define

$$C_c = \{(c_1, \ldots, c_{n-1}): (c_1, c_2, \ldots, c_{n-1}, c) \in C\}, \tag{2.5}$$

$$J(C) = \{c: C_c \neq \emptyset\}. \tag{2.6}$$

and

$$L(C) = \max\Big\{|D|: D \subset J(C), \bigcap_{c \in D} C_c \neq \varnothing\Big\}. \tag{2.7}$$

LEMMA 1. *For* $(A, B) \in \mathscr{P}(f, \mathscr{R}, n)$, $L(A)|J(B)| \leq M(f, \mathscr{R}, 1)$.

Proof. It suffices to show that for every $D \subset J(A)$ with $\bigcap_{a \in D} A_a \neq \varnothing$, necessarily $(D, J(B)) \in \mathscr{P}(f, \mathscr{R}, 1)$.

To see this choose any $a, a' \in D$ and any $b, b' \in J(B)$ and use the fact that by our assumptions there are a^{n-1}, b^{n-1}, and b'^{n-1} such that $a^{n-1}a, a^{n-1}a' \in A$ and $b^{n-1}b, b'^{n-1}b' \in B$. Now just verify

$$\begin{aligned}\mathscr{R} \ni\ & f_n(a^{n-1}a, b^{n-1}b) - f_n(a^{n-1}a, b'^{n-1}b') \\ & + f_n(a^{n-1}a', b'^{n-1}b') - f_n(a^{n-1}a', b^{n-1}b) \\ =\ & f(a, b) - f(a, b') + f(a', b') - f(a', b),\end{aligned}$$

since

$$\begin{aligned}&\big[f_{n-1}(a^{n-1}, b^{n-1}) - f_{n-1}(a^{n-1}, b'^{n-1}) \\ &+ f_{n-1}(a^{n-1}, b'^{n-1}) - f_{n-1}(a^{n-1}, b^{n-1})\big] = 0.\end{aligned}$$

LEMMA 2. *If* $(A, B) \in \mathscr{P}(f, \mathscr{R}, n)$, *then* $(\bigcup_{d \in J(A)} A_d, B_d) \in \mathscr{P}(f, \mathscr{R}, n-1)$ *for all* $b \in J(B)$.

Proof. For $a^{n-1}, a'^{n-1} \in \bigcup_{d \in J(A)} A_d$ there exist $a, a' \in J(A)$ such that $a^{n-1} \in A_a$ and $a'^{n-1} \in A_{a'}$. Now for any $b^{n-1}, b'^{n-1} \in B_b$,

$$\begin{aligned}\mathscr{R} \ni\ & f_n(a^{n-1}a, b^{n-1}b) - f_n(a^{n-1}a, b'^{n-1}b) \\ & + f_n(a'^{n-1}a', b'^{n-1}b) - f_n(a'^{n-1}a', b^{n-1}b) \\ =\ & f_{n-1}(a^{n-1}, b^{n-1}) - f_{n-1}(a^{n-1}, b'^{n-1}) \\ & + f_{n-1}(a'^{n-1}, b'^{n-1}) - f_{n-1}(a'^{n-1}, b^{n-1}),\end{aligned}$$

since $[f(a, b) - f(a, b) + f(a', b) - f(a', b)] = 0$.

Proof of Theorem 1. Obviously, if for $(A, B) \in \mathscr{P}^*(f, \{0\}, 1)$ we have $|A|\,|B| = M(f, \mathscr{R}, 1)$, then $(\prod_1^n A, \prod_1^n B) \in \mathscr{P}(f, \{0\}, n)$ and therefore $M(f, \mathscr{R}, n) \geq (|A|\,|B|)^n = M(f, \mathscr{R}, 1)^n$

We now prove (i) by induction on n. (For $n = 1$ nothing needs to be proved.) For $(A, B) \in \mathscr{P}(f, \mathscr{R}, n)$ we have

$$|A||B| = \sum_{a \in J(A)} |A_a| \cdot \sum_{b \in J(B)} |B_b|$$

$$\leq L(A) \Big| \bigcup_{a \in J(A)} A_a \Big| |J(B)| \max_{b \in J(B)} |B_b|$$

$$\leq M(f, \mathscr{R}, 1) \Big| \bigcup_{a \in J(A)} A_a \Big| \max_{b \in J(B)} |B_b| \qquad \text{(by Lemma 1)}.$$

The result $|A||B| \leq M(f, \mathscr{R}, 1)^n$ now follows from Lemma 2 and the induction hypothesis.

Remark. T. Scheuer kindly pointed out to us that essentially the same proof gives a more general result: $\mathbb{Z}$ can be replaced by any abelian group and for $\{0\}$ one can allow any subgroup of $\mathbb{Z}$.

3. Old and New Results Implied by Theorem 1

We show first that the original 4-words inequality of [4] is implied by Theorem 1.

Corollary 1. *If $A, B \subset \{1, 2, \ldots, \alpha\}^n$ and for the Hamming distance function d_H*

$$d_H(a^n, b^n) - d_H(a^n, b'^n) + d_H(a'^n, b'^n) - d_H(a'^n, b^n) \neq 1, 2 \quad (3.1)$$

for all $a^n, a'^n \in A$ and all $b^n, b'^n \in B$, then

$$|A||B| \leq \alpha^{*n}, \qquad \text{where } \alpha^* = \begin{cases} \alpha & \text{for } \alpha = 2, 3, 4 \\ \lceil \alpha/2 \rceil \lfloor \alpha/2 \rfloor & \text{for } \alpha \geq 4. \end{cases}$$

Furthermore, the bound is best possible.

Proof. Just notice that condition (3.1) says that $(A, B) \in \mathscr{P}(d_H, \mathbb{Z} - \{1, 2\}, n)$ and that $\mathscr{P}(d_H, \mathbb{Z} - \{1, 2\}, 1) = \mathscr{P}(d_H, \{0\}, 1)$. Therefore by Theorem 1, $M(d_H, \mathbb{Z} - \{1, 2\}, n) = M(d_H, \{0\}, 1)^n$. Finally the equality $M(d_H, \{0\}, 1) = \alpha^*$ is readily verified.

Next we derive the 4-words inequality of [7] concerning the Lee metric d_1.

COROLLARY 2. *If* $A, B \subset \{0,1,2,\ldots,\alpha-1\}^n$ *and*

$$d_L(a^n,b^n) - d_L(a^n,b'^n) + d_L(a'^n,b'^n) - d_L(a'^n,b^n) \neq 1,2,\ldots,\alpha \text{ for all } a^n, a'^n \in A \text{ and all } b^n, b'^n \in B, \text{ then } |A|\,|B| \leq (\max[\alpha, (\lfloor \alpha/4 \rfloor + 1)(\lceil \lfloor \alpha/2 \rfloor/2 \rceil + 1)])^n \text{ and this bound is best possible.} \tag{3.2}$$

Proof. Here condition (3.2) says that $(A, B) \in \mathscr{P}(d_L, \mathbb{Z} - \{1,2,\ldots,\alpha\}, n)$. For $n = 1$ thus necessarily $\mathscr{P}(d_L, \mathbb{Z} - \{1,2,\ldots,\alpha\}, 1) = \mathscr{P}(d_L, \{0\}, 1)$ and by Theorem 1, $M(d_L, \mathbb{Z} - \{1,2,\ldots,\alpha\}, n) = M(d_L, \{0\}, 1)^n$. It can be calculated that

$$M(d_L,\{0\},1) = \begin{cases} \alpha & \text{for } \alpha = 2,3,4,5,6,7,9 \\ (\lfloor \alpha/4 \rfloor + 1)(\lceil \lfloor \alpha/2 \rfloor/2 \rceil + 1) & \text{otherwise.} \end{cases} \tag{3.3}$$

For $\alpha = 2$ one has $d_L = d_H$. Another metric coinciding with d_H for $\alpha = 2$ is the Taxi metric (L_1-metric in analysis), which for $a^n, b^n \in \{0,1,\ldots,\alpha-1\}^n$ is defined by

$$d_T(a^n,b^n) = \sum_{i=1}^{n} |a_i - b_i|. \tag{3.4}$$

The following result is new.

COROLLARY 3. *If* $A, B \subset \{0,1,\ldots,\alpha-1\}^n$ *and*

$$d_T(a^n,b^n) - d_T(a^n,b'^n) + d_T(a'^n,b'^n) - d_T(a'^n,b^n) \neq 1,2,\ldots,2\alpha \text{ for all } a^n, a'^n \in A \text{ and all } b^n, b'^n \in B, \text{ then } |A|\,|B| \leq (\max[\alpha, (\lfloor \alpha/2 \rfloor + 1)\lceil \alpha/2 \rceil])^n \text{ and this bound is best possible.} \tag{3.5}$$

Proof. Condition (3.5) says that $(A, B) \in \mathscr{P}(d_T, \mathbb{Z} - \{1,2,\ldots,2\alpha\}, n)$. Therefore $\mathscr{P}(d_T, \mathbb{Z} - \{1,2,\ldots,2\alpha\}, 1) = \mathscr{P}(d_T, \{0\}, 1)$ and by Theorem 1, $M(d_T, \mathbb{Z} - \{1,2,\ldots,2\alpha\}, n) = M(d_T, \{0\}, 1)^n$. It remains to be seen that

$$M(d_T,\{0\},1) = \max[\alpha, (\lfloor \alpha/2 \rfloor + 1)\lceil \alpha/2 \rceil]. \tag{3.6}$$

By choosing $(A, B) = (\{0\}, \{0,1,\ldots,\alpha-1\})$ or $(A, B) = (\{0,1,\ldots,\lfloor \alpha/2 \rfloor\}, \{\lfloor \alpha/2 \rfloor,\ldots,\alpha-1\})$ in $\mathscr{P}(d_T, \{0\}, 1)$ we get

$$M(d_T,\{0\},1) \geq \max[\alpha, (\lfloor \alpha/2 \rfloor + 1)\lceil \alpha/2 \rceil].$$

Now we prove the reverse inequality. Clearly, w.l.o.g. we can assume that $|A| \geq |B|$. Furthermore, we are done if we can show that

$$a \leq b \ (\text{or } a \geq b) \qquad \text{for all } a \in A, b \in B. \tag{3.7}$$

Assume to the contrary that there are $a \in A$; $b, b' \in B$ with $b' < a < b$. Then $b' < a' < b$ must hold for all $a' \in A$, because otherwise $d_T(a', b) - d_T(a', b') = b - b' > d_T(a, b) - d_T(a, b')$, which contradicts $(A, B) \in \mathscr{P}(d_T, \{0\}, 1)$.

However, if now $b' < a' < b$ for all $a' \in A$, then

$$\begin{aligned} a' &= a' - \tfrac{1}{2}[d_T(a, b) - d_T(a, b') + d_T(a', b') - d_T(a', b)] \\ &= a' - \tfrac{1}{2}[(b - a) - (a - b') + (a' - b') - (b - a')] \\ &= a' - \tfrac{1}{2}[2a' - 2a] = a, \end{aligned}$$

and hence $|A| = 1 < 2 \leq |B|$, which contradicts $|A| \geq |B|$.

As last examples, we consider the Boolean functions union "$\vee$" and intersection "$\wedge$." Here it is assumed that $\mathscr{X} = \mathscr{Y} = \{0, 1\}$. In earlier work [1] the Boolean function symmetric difference "Δ" was studied.

Actually, we are interested in cardinalities of unions and intersections, that is, the functions μ_n, λ_n: $\mathscr{X}^n \times \mathscr{Y}^n \to \mathbb{N}$, where

$$\mu_n(a^n, b^n) = \sum_{t=1}^{n} a_t \vee b_t, \tag{3.8}$$

$$\lambda_n(a^n, b^n) = \sum_{t=1}^{n} a_t \wedge b_t. \tag{3.9}$$

Not only do we consider $\mathscr{P}(\mu, \mathscr{R}, n)$ and $M(\mu, \mathscr{R}, n)$, but also the following sets and functions

$$\begin{aligned} \mathscr{P}^*(\mu, \delta, n) = \{(A, B)\colon A, B \subset \mathscr{X}^n \text{ with } \mu_n(a^n, b^n) \\ = \delta \text{ for all } a^n \in A, b^n \in B\}, \end{aligned} \tag{3.10}$$

$$M^*(\mu, \delta, n) = \max\{|A||B|\colon (A, B) \in \mathscr{P}^*(\mu, \delta, n)\}, \tag{3.11}$$

$$M^*(\mu, n) = \max_{0 \leq \delta \leq n} M^*(\mu, \delta, n). \tag{3.12}$$

Analogously $\mathscr{P}^*(\lambda, \delta, n)$, $M^*(\lambda, \delta, n)$, and $M^*(\lambda, n)$ are defined. Results for μ can be transformed into results for λ by complementation. Thus $M^*(\mu, \delta, n) = M^*(\lambda, n - \delta, n)$, etc. Now obviously $M^*(\mu, \delta, n) \leq M(\mu, \{0\}, n)$ and, by Theorem 1, $M(\mu, \mathbb{Z} \setminus \{1, 2\}, n) = M(\mu, \{0\}, n) = M(\mu, \{0\})^n = 2^n$. On the other hand, $(A, B) = (\{11 \ldots 1\}, \mathscr{X}^n) \in \mathscr{P}^*(\mu, n, n)$ and $|A||B| = 2^n$.

COROLLARY 4. $M(\mu, \mathbb{Z} - \{1,2\}, n) = M^*(\mu, n) = 2^n$.

Remark 1. There is another way to show that $M^*(\lambda, n) = M^*(\mu, n) = 2^n$. By (3.9) we know that the inner product (a^n, b^n) equals $\lambda_n(a^n, b^n)$. Furthermore, if $(a^n, b^n) = \text{const.}$ for all $a^n \in A$, $b^n \in B$, then also $a^n \circ b^n = \sum_{t=1}^n a_t \wedge b_t$ with addition understood mod 2 (that is, the inner product of a^n, b^n as vectors of a linear space over GF(2)) is constant. From the fact $(x^n \oplus a^n) \circ b^n = (x^n \circ b^n) \oplus (a^n \circ b^n)$, we conclude that $a'^n \oplus A$ and B are orthogonal, if $a'^n \in A$. Thus, $\dim A + \dim B \leq n$ and $|A||B| \leq 2^n$. Actually, if $a^n \circ b^n = 1$, for all $a^n \in A$, $b^n \in B$, then it can be shown that $|A||B| \leq 2^{n-1}$.

4. AN EXACT RESULT FOR "ONE-SIDED CONSTANT" UNIONS CARDINALITIES

Most results of the remainder of this paper originated in attempts to prove the

Conjecture. $M^*(\mu, \delta, n) = \max_{0 \leq m \leq \delta} 2^m \binom{n-m}{\delta - m}$.

They have led to a positive answer at least in an asymptotic sense (Section 5). One of the ideas of [4] was to generalize the concept of a constant distance code pair to that of a one-sided constant distance code pair, which is better suited for inductive arguments. Further generalization led to notions of 4-word properties. The same idea also turns out to be useful for the function μ. Actually we introduce a refined, that is, quantitative, notion of 1-sidedness,

$$\mathscr{P}(\mu, \leq \delta, n) = \{(A, B): A, B \subset \mathscr{X}^n; |a \cup b| = |a' \cup b| \leq \delta \text{ for all } a, a' \in A, b \in B\} \quad (4.1)$$

$$\mathbf{M}(\mu, \leq \delta, n) = \max\{|A||B|: (A, B) \in \mathscr{P}(\mu, \leq \delta, n)\}. \quad (4.2)$$

THEOREM 2. $\mathbf{M}(\mu, \leq \delta, n) = \sum_{k=0}^{\delta} \binom{n}{k}$.

Proof. Since $(\{00\ldots0\}, \{b^n: \sum_{t=1}^n b_t \leq \delta\}) \in \mathscr{P}(\mu, \leq \delta, n)$, we have $\mathbf{M}(\mu, \leq \delta, n) \geq \sum_{k=0}^{\delta} \binom{n}{k}$.

The opposite inequality can be proved by induction on n. The cases $n = 1, 2$ are done by inspection. Since $(A, B) \in \mathscr{P}(\mu, \leq \delta, n)$ implies

$(A, B) \in \mathscr{P}(\mu, \{0\}, n)$, by Corollary 4 also for $\delta = n$, $|A||B| \le 2^n = \sum_{k=0}^{\delta}\binom{n}{k}$.

For the induction we use the familiar definition

$$C_c^t = \{(c_1, \ldots, c_{t-1}, c_{t+1}, \ldots, c_n) : (c_1, \ldots, c_{t-1}, c, c_{t+1}, \ldots, c_n) \in C\}, \quad \text{if } C \subset \mathscr{X}^n. \tag{4.3}$$

$n - 1 \to n$. When $\delta < n$, then $A \subset \mathscr{X}^n \setminus \{(1, 1, \ldots, 1)\}$. Thus for some t, $A_0^t \neq \varnothing$. We need the definitions

$$B(k) = \{b \in B : |a \cup b| = k \text{ for all } a \in A\} \tag{4.4}$$

$$B_\epsilon^t(k) = B(k) \cap B_\epsilon^t. \tag{4.5}$$

Case $A_1^t = \varnothing$.

$$\begin{aligned} |A||B| &= |A| \sum_{k=0}^{\delta} |B(k)| \\ &= |A_0^t| \sum_{k=0}^{\delta} |B_0^t(k)| + |A_0^t| \sum_{k=0}^{\delta} |B_1^t(k)| \\ &\le \sum_{k=0}^{\delta} \binom{n-1}{k} + \sum_{k=0}^{\delta-1} \binom{n-1}{k} \qquad \text{(by induction hypothesis)} \\ &= \sum_{k=0}^{\delta} \binom{n}{k}. \end{aligned}$$

Case $A_1^t \neq \varnothing$. The relations $A_0^t \neq \varnothing$ and $A_1^t \neq \varnothing$ have the following consequences:

$$B_1^t(0) = \varnothing, \qquad \text{because } A \neq \varnothing. \tag{4.6}$$

$$B_0^t(k) \cap B_1^t(l) = \varnothing \qquad \text{for } k \le l, \text{ because } A_0^t \neq \varnothing. \tag{4.7}$$

$$B_0^t(0) = \varnothing, \qquad \text{because } A_1^t \neq \varnothing. \tag{4.8}$$

$$\text{For } b \in B_0^t(k) \text{ and all } a \in A_0^t, |a \cup b| = k. \tag{4.9}$$

$$\text{For } b \in B_0^t(k) \cup B_1^t(k) \text{ and all } a \in A_1^t, |a \cup b| = k - 1. \tag{4.10}$$

These facts ensure the following chain of equalities and inequalities:

$$
\begin{aligned}
|A|\,|B| &= |A| \sum_{k=0}^{\delta} |B(k)| \\
&= \left(|A_0'| + |A_1'|\right) \sum_{k=0}^{\delta} \left(|B_0'(k)| + |B_1'(k)|\right) \\
&= |A_0'|\left\{ \sum_{k=0}^{\delta} |B_0'(k)| + \sum_{k=0}^{\delta-1} |B_1'(k+1)| \right\} \\
&\quad + |A_1'|\left\{ \sum_{k=0}^{\delta} |B_0'(k)| + \sum_{k=1}^{\delta} |B_1'(k)| \right\} \qquad \text{by (4.6), (4.8)} \\
&= |A_0'|\left\{ \sum_{k=0}^{\delta-1} |B_0'(k) \cup B_1'(k+1)| + |B_0'(\delta)| \right\} \\
&\quad + |A_1'| \sum_{k=0}^{\delta} |B_0'(k) \cup B_1'(k)| \qquad \text{(by (4.7))} \\
&\le \sum_{k=0}^{\delta} \binom{n-1}{k} + \sum_{k=0}^{\delta-1} \binom{n-1}{k} \\
&\qquad \text{(by induction hypothesis and (4.9), (4.10))} \\
&= \sum_{k=0}^{\delta} \binom{n}{k}.
\end{aligned}
$$

5. The Asymptotic Behaviour of $M^*(\mu, \delta, n)$

We assume now that $(\delta(n))_{n=1}^{\infty}$ is a sequence of non-negative integers with

$$\lim_{n \to \infty} \delta(n) n^{-1} = \varepsilon \in [0,1]. \tag{5.1}$$

The construction behind our conjecture in Section 4 is

$$
\begin{aligned}
A &= \{ a^n : a_t = 1 \text{ exactly when } 1 \le t \le m \} \\
B &= \left\{ b^n \in \mathcal{X}^n : \sum_{t=m+1}^{n} b_t = \delta - m \right\}.
\end{aligned} \tag{5.2}
$$

Now choose

$$m = \begin{cases} 2\delta(n) - n & \text{when } 2\delta(n) > n \\ 0 & \text{when } 2(n) \le n. \end{cases} \tag{5.3}$$

Then

$$\frac{1}{n}\log|A||B| = \begin{cases} \frac{1}{n}\log 2^{2\delta(n)-n}\binom{2n-2\delta(n)}{n-\delta(n)} & \text{if } 2\delta(n) > n, \\ \frac{1}{n}\log\binom{n}{\delta(n)} & \text{if } 2\delta(n) \le n, \end{cases}$$

$$\to \begin{cases} 1 & \text{if } \varepsilon \ge \frac{1}{2} \\ h(\varepsilon) & \text{if } \varepsilon < \frac{1}{2} \text{ as } n \to \infty, \end{cases}$$

where $h(\varepsilon) = -\varepsilon\log\varepsilon - (1-\varepsilon)\log(1-\varepsilon)$, is the binary entropy function. This result is asymptotically best possible.

THEOREM 3. *If* (5.1) *holds, then*

$$\lim_{n\to\infty}\frac{1}{n}\log M^*(\mu, \delta(n), n) = \begin{cases} h(\varepsilon) & \text{for } \varepsilon < \frac{1}{2} \\ 1 & \text{for } \varepsilon > \frac{1}{2}. \end{cases}$$

Proof. It remains to be seen that the quantities to the right side cannot be superceded. Since $M^*(\mu, \delta(n), n) \le \mathbf{M}(\mu, \le \delta(n), n)$ and since by Theorem 2,

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbf{M}(\mu, \le \delta(n), n) = \max_{0\le p\le\varepsilon} h(p) = \begin{cases} h(\varepsilon) & \text{for } \varepsilon < \frac{1}{2}, \\ 1 & \text{for } \varepsilon \ge \frac{1}{2}, \end{cases}$$

this is the case.

6. REPLACING CARDINALITIES BY PRODUCT MEASURES

In all results of the preceding sections the sets where measured by cardinality. More generally, now we use a measure P on $\mathscr{X}$ and a measure Q on $\mathscr{Y}$. Sets in $\mathscr{X}^n$ (resp. $\mathscr{Y}^n$) are measured by the product measure $P^n = \prod_1^n * P$ (resp. $Q^n = \prod_1^n * Q$).

A first result in this setting is a generalization of Theorem 1. Let us define

$$M(f, \mathscr{R}, n, P, Q) = \max\{P^n(A)Q^n(B) | (A, B) \in \mathscr{P}(f, \mathscr{R}, n)\}. \tag{6.1}$$

Instead of (2.7) we need now

$$L(A, P) = \max\Big\{P(D): D \subset J(A), \bigcap_{d\in D} A_d \neq \varnothing\Big\}. \tag{6.2}$$

Inspection of the proof of Lemma 1 shows that we have now

LEMMA 1′. *For* $(A, B) \in \mathscr{P}(f, \mathscr{R}, n)$,

$$L(A, P)Q(J(B)) \leq M(f, \mathscr{R}, 1, P, Q).$$

Lemma 1′ is used in the proof for the following theorem, which generalizes Theorem 1.

THEOREM 1′. *For any* $\mathscr{R} \subset \mathbb{Z}$,

$$M(f, \mathscr{R}, n, P, Q) \leq M(f, \mathscr{R}, 1, P, Q)^n. \tag{i}$$

Furthermore, if $0 \in \mathscr{R}$ *and* $M(f, \{0\}, 1, P, Q) = M(f, \mathscr{R}, 1, P, Q)$, *then equality holds in* (i).

Proof. For $(A, B) \in \mathscr{P}(f, \mathscr{R}, n)$, we have

$$\begin{aligned}
&P^n(A)Q^n(B)\\
&\quad = \sum_{a\in J(A)} P(a)P^{n-1}(A_a) \cdot \sum_{b\in J(B)} Q(b)Q^{n-1}(B_b)\\
&\quad \leq L(A, P)P^{n-1}\Big(\bigcup_{a\in J(A)} A_a\Big)Q(J(B)) \max_{b\in J(B)} Q^{n-1}(B_b)\\
&\quad \leq L(A, B)Q(J(B))M(f, \mathscr{R}, n-1, P, Q) \qquad \text{(by Lemma 2)}\\
&\quad \leq M(f, \mathscr{R}, 1, P, Q)M(f, \mathscr{R}, n-1, P, Q) \qquad \text{(by Lemma 1′)}\\
&\quad \leq M(f, \mathscr{R}, 1, P, Q)^n \qquad \text{(by induction hypothesis).}
\end{aligned}$$

The last statement of the theorem follows as before.

Our next result is an extension of Theorem 2 (Section 4) in *two directions*. Cardinality is replaced by suitable product measures and the condition "one-sided constant" is *substituted for by a more general 4-words property*.

THEOREM 2′. *Let* $\mathscr{X} = \mathscr{Y} = \{0, 1\}$ *and let* P^n *and* Q^n *be product measures on* $\{0, 1\}^n$ *such that*

(a) $P(0) \geq P(1)$ *and* $Q(0) \geq Q(1)$. *If an* $(A, B) \in \mathscr{P}(\mu, \mathbb{Z} - \{1\}, n)$ *satisfies*

(b) $\mu_n(a^n, b^n) \leq \delta$ *for all* $a^n \in A$, $b^n \in B$ *then*

(c) $P^n(A)Q^n(B) \leq \max\{P(0)^n\Sigma_{k=0}^{\delta}\binom{n}{k}Q(1)^kQ(0)^{n-k}$, $Q(0)^n\Sigma_{k=0}^{\delta}\binom{n}{k}P(1)^kP(0)^{n-k}\}$ *and this bound is best possible.*

Proof. The last statement is readily verified. For $(A, B) = (\{(0, \ldots, 0)\}, \{y^n: \Sigma_{t=1}^n y_t \leq \delta\}) \in \mathscr{P}(\mu, \mathbb{Z} - \{1\}, n)$, (b) holds and $P^n(A)Q^n(B) = P(0)^n \Sigma_{k=0}^{\delta} \binom{n}{k} Q(1)^k Q(0)^{n-k}$.

If this is the smaller of the two values, choose the pair (B, A). We now prove (b) by induction on n. We can enforce, by exchanging the roles of P and Q if necessary,

$$P(0)Q(1) \geq P(1)Q(0). \tag{6.3}$$

This implies

$$\begin{aligned} P(0)^n Q(1)^k Q(0)^{n-k} &= (P(0)Q(1))^k (P(0)Q(0))^{n-k} \\ &\geq (P(1)Q(0))^k (P(0)Q(0))^{n-k} \\ &= Q(0)^n P(1)^k P(0)^{n-k} \end{aligned}$$

and we therefore have to prove for $(A, B) \in \mathscr{P}(\mu_n, \mathbb{Z} - \{1\}, n)$ satisfying (b),

$$P^n(A)Q^n(B) \leq P(0)^n \sum_{k=0}^{\delta} \binom{n}{k} Q(1)^k Q(0)^{n-k}. \tag{6.4}$$

$n = 1, \delta = 0$. The only choice is $A = B = \{0\}$ and here (6.4) holds.

$n = 1, \delta = 1$. The case $A = B = \{0, 1\}$ does not arise and otherwise $P(A)Q(B) \leq P(0)(Q(0) + Q(1))$ by (a) and (6.3).

$n - 1 \to n$. If $\delta = n$, then by Theorem 1′ and (6.3),

$$\begin{aligned} P^n(A)Q^n(B) &\leq (P(0)(Q(0) + Q(1))^n \\ &= P(0)^n \sum_{k=0}^{n} \binom{n}{k} Q(1)^k Q(0)^{n-k}; \end{aligned}$$

that is, (6.4) holds.

If $\delta < n$ then B cannot contain the "all 1 sequence" and therefore $B_0^t \neq \varnothing$ for some $t \in \{1, 2, \ldots, n\}$. In case also $B_1^t \neq \varnothing$, we claim that

$$A_0^t \cap A_1^t = \varnothing, \tag{6.5}$$

because otherwise there are $b^{n-1} \in B_0^t$, $b^{n-1} \in B_1^t$ and $a^{n-1} \in A_0^t \cap A_1^t$, which satisfy

$$\begin{aligned} 1 \neq \mu_n(1a^{n-1}, 0b^{n-1}) - \mu_n(1a^{n-1}, 1b^{n-1}) \\ + \mu_n(0a^{n-1}, 1b^{n-1}) - \mu_n(0a^{n-1}, 0b^{n-1}) = 1, \end{aligned}$$

a contradiction.

Therefore $P^{n-1}(A_0') + P^{n-1}(A_1') = P^{n-1}(A_0' \cup A_1')$ and the inequalities in (a) yield

$$\begin{aligned} P^n(A)Q^n(B) &= \big(P(0)P^{n-1}(A_0') + P(1)P^{n-1}(A_1')\big) \\ &\quad \times \big(Q(0)Q^{n-1}(B_0') + Q(1)Q^{n-1}(B_1')\big) \\ &\le P(0)P^{n-1}(A_0' \cup A_1')Q(0)\big[Q^{n-1}(B_0') + Q^{n-1}(B_1')\big]. \end{aligned}$$

By Lemma 2 and the induction hypothesis, therefore,

$$\begin{aligned} P^n(A)Q^n(B) &\le P(0)Q(0)\Bigg[P^{n-1}(0)\sum_{k=0}^{\delta}\binom{n-1}{k}Q(1)^kQ(0)^{n-k-1} \\ &\qquad + P^{n-1}(0)\sum_{k=0}^{\delta-1}\binom{n-1}{k}Q(1)^kQ(0)^{n-k-1}\Bigg] \\ &= P(0)^n\Bigg[\sum_{k=0}^{\delta}\binom{n-1}{k}Q(1)^kQ(0)^{n-k} \\ &\qquad + \sum_{k=0}^{\delta-1}\binom{n-1}{k}Q(1)^kQ(0)^{n-k}\Bigg] \\ &= P(0)^n\sum_{k=0}^{\delta}\binom{n}{k}Q(1)^kQ(0)^{n-k}. \end{aligned}$$

We are left with the cases $B_0' \neq \varnothing$ and $B_1' = \varnothing$;

$$\begin{aligned} P^n(A)Q^n(B) &= \big(P(0)P^{n-1}(A_0') + P(1)P^{n-1}(A_1')\big)Q(0)Q^{n-1}(B_0') \\ &\le P(0)Q(0)P^{n-1}(A_0')Q^{n-1}(B_0') \\ &\quad + P(0)Q(0)P^{n-1}(A_1')Q^{n-1}(B_0') \end{aligned}$$

and, by Lemma 2 and the induction hypothesis,

$$\begin{aligned} P^n(A)Q^n(B) &\le P(0)^n\Bigg[Q(0)\sum_{k=0}^{\delta}\binom{n-1}{k}Q(1)^kQ(0)^{n-k-1} \\ &\qquad + Q(0)\sum_{k=0}^{\delta-1}\binom{n-1}{k}Q(1)^kQ(0)^{n-k-1}\Bigg] \\ &= P(0)^n\sum_{k=0}^{\delta}\binom{n}{k}Q(1)^kQ(0)^{n-k}. \end{aligned}$$

7. From the Union Function μ_n to General "Sum-Type" Functions

The Hamming distance and μ_n are of sum type. Replacing μ in the definitions (3.10) to (3.12) by an arbitrary f: $\mathcal{X} \times \mathcal{Y} \to \mathbb{Z}$ we get the set $P^*(f, \delta, n)$ and the numbers $M^*(f, \delta, n)$ and $M^*(f, n)$. We shall establish relations to $M(f, \{0\}, n)$, which was defined in Section 2. For this the following result is basic.

Decomposition Lemma. *$(A, B) \in \mathcal{P}(f, \{0\}, 1)$ if and only if there exist functions $\varphi_{A,B}$: $A \to \mathbb{Z}$, $\psi_{A,B}$: $B \to \mathbb{Z}$ such that*

$$f(a, b) = \varphi_{A,B}(a) + \psi_{A,B}(b) \quad \text{for all } (a, b) \in A \times B. \tag{7.1}$$

Proof. Clearly, if f satisfies (7.1), then

$$\begin{aligned} &f(a, b) - f(a', b) + f(a', b') - f(a, b') \\ &\quad = \varphi_{A,B}(a) - \varphi_{A,B}(a') + \psi_{A,B}(a') - \psi_{A,B}(a) = 0 \end{aligned}$$

and therefore $(A, B) \in \mathcal{P}(f, \{0\}, 1)$.

Conversely, for $(A, B) \in \mathcal{P}(f, \{0\}, 1)$ with $A = \{a_0, \ldots, a_l\}$ and $B = \{b_0, \ldots, b_m\}$ the functions

$$\varphi_{A,B}(a) = f(a, b_0) \quad \text{and} \quad \psi_{A,B}(b) = f(a_0, b) - f(a_0, b_0) \tag{7.2}$$

satisfy $\varphi_{A,B}(a) + \psi_{A,B}(b) = f(a, b_0) + f(a_0, b) - f(a_0, b_0)$ and, by the 4-word property, thus (7.1).

Theorem 4. (a) $M(f, \{0\}, n) = M(f, \{0\}, 1)^n$
(b) $\lim_{n \to \infty} (1/n) \log M^*(f, n) = M(f, \{0\}, 1)$.

Proof. (a) is an immediate consequence of Theorem 1. Also from the definitions,

$$M^*(f, n) \le M(f, \{0\}, n). \tag{7.3}$$

The issue is to show that asymptotically the reverse inequality also holds.

For this suppose that (A, B) achieves $M(f, \{0\}, 1)$ and that $\varphi_{A,B}$ and $\psi_{A,B}$ are defined as in the Decomposition Lemma, then for $(A^n, B^n) = (\prod_1^n A, \prod_1^n B)$,

$$|A^n||B^n| = M(f, \{0\}, 1)^n \tag{7.4}$$

and

$$f(x^n, y^n) = \sum_{t=1}^{n} \varphi_{A,B}(x_t) + \psi_{A,B}(y_t) \qquad \text{for } (x^n, y^n) \in A^n \times B^n. \quad (7.5)$$

With the definitions

$$\varphi_{A^n,B^n}(x^n) = \sum_{t=1}^{n} \varphi_{A,B}(x_t), \qquad \psi_{A^n,B^n}(y^n) = \sum_{t=1}^{n} \psi_{A,B}(y_t), \quad (7.6)$$

we have, therefore,

$$f_n(x^n, y^n) = \varphi_{A^n,B^n}(x^n) + \psi_{A^n,B^n}(y^n) \qquad \text{for } (x^n, y^n) \in A^n \times B^n. \quad (7.7)$$

The key observation now is that φ_{A^n,B^n} and ψ_{A^n,B^n} have as sum-type functions a rather small range of values on A^n (resp. B^n) *and* that by (7.7) for any $u, v \in \mathbb{Z}$ f_n is constant on $U \times V$, where $U = \varphi^{-1}_{A^n,B^n}(u)$ and $V = \psi^{-1}_{A^n,B^n}(v)$. The formal argument follows.

Defining

$$\Phi = \max_{x \in \mathcal{X}} \varphi_{AB}(x) - \min_{x \in \mathcal{X}} \varphi_{AB}(x) + 1,$$
$$\psi = \max_{y \in \mathcal{Y}} \psi_{AB}(y) - \min_{y \in \mathcal{Y}} \psi_{AB}(y) + 1, \quad (7.8)$$

and defining by $\|g\|$ the cardinality of the range of a function g, we derive with (7.6)

$$\|\varphi_{A^nB^n}\| \le n\Phi, \qquad \|\psi_{A^nB^n}\| \le n\psi. \quad (7.9)$$

We conclude now that

$$M^*(f, n) \ge \frac{|A^n||B^n|}{n^2\Phi\psi} = \frac{M(f, \{0\}, 1)^n}{n^2\Phi\psi}, \quad (7.10)$$

which together with (7.3) implies (b).

8. Applications to Two-Way Communication Complexity

A. Preliminaries

After Abelson had raised the issue of information transfer in distributed computations [9], Yao did his pioneering work on two-way communication complexity [5]. His success is mainly due to his limitation to functions

$f\colon \mathcal{X}\times\mathcal{Y}\to\mathcal{Z}$ with $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ finite, which made a combinatorial treatment possible. A natural improvement of Yao's model [10] led to a very smooth form of Yao's lower bound for the two-way complexity $C(f; 1 \leftrightarrow 2)$, which we now state without proof:

$$S \times T (S \subset \mathcal{X}, T \subset \mathcal{Y}) \text{ is called } f\text{-monochromatic,}$$
$$\text{if } f \text{ is constant on } S \times T.$$

A k-decomposition of f is a partition $S = \{S_1 \times T_1, \ldots, S_k \times T_k\}$ of $\mathcal{X}\times\mathcal{Y}$ into f-monochromatic rectangles.

For the decomposition number

$$D(f) \triangleq \min\{k\colon \text{exists a } k\text{-decomposition of } f\}. \tag{8.1}$$

Yao's inequality (in the improved form of [10]) states

$$C(f; 1 \leftrightarrow 2) \geq \log_2 D(f). \tag{8.2}$$

We have not yet defined $C(f; 1 \leftrightarrow 2)$. It is actually a quantity which can be understood without any reference to computing in the context of an abstract multi-user source coding theory (see [1]).

The specifica here are:

(1) No probabilistic assumptions on the source $(\mathcal{X}, \mathcal{Y}, f)$

(2) Correct decoding for all source outputs.

The communication model is: $\mathcal{X}$ outputs x and $\mathcal{Y}$ outputs y. A person $P_{\mathcal{X}}$ observes x and another person $P_{\mathcal{Y}}$ observes y. They can transmit messages to each other alternately over a binary channel with zero error probability. Theor goal is to find the value $f(x, y)$ with minimal worst case transmission time. This quantity is denoted by $C(f; 1 \leftrightarrow 2)$.

Similar to classical source coding there is a multitude of other communication models one might consider. There is a trivial general bound on $C(f; 1 \leftrightarrow 2)$.

Suppose that $P_{\mathcal{X}}$ transmits the output x of $\mathcal{X}$ of $P_{\mathcal{Y}}$, who in the knowledge of x and y calculates $f(x, y)$ and transmits this value to $P_{\mathcal{X}}$, then obviously $\lceil \log_2 |\mathcal{X}| \rceil + \lceil \log_2 |\mathcal{Z}| \rceil$ bits suffice, i.e.,

$$C(f; 1 \leftrightarrow 2) \leq \lceil \log_2 |\mathcal{X}| \rceil + \lceil \log_2 |\mathcal{Z}| \rceil. \tag{8.3}$$

There is also a general, but naive, lower bound on $D(f)$.

Denoting the size of the largest monochromatic rectangle of f by $M(f)$, we clearly have $D(f) \geq |\mathcal{X}||\mathcal{Y}|M(f)^{-1}$ and thus by (8.2),

$$C(f; 1 \leftrightarrow 2) \geq \left\lceil \log_2 |\mathcal{X}||\mathcal{Y}|M(f)^{-1} \right\rceil. \tag{8.4}$$

In some cases (8.3) or (and) (8.4) give good estimates on $C(f; 1 \leftrightarrow 2)$.

B. A New Lower Bound

The use of the 4-words inequality (Theorem 1) for complexity problems is due to its property to relate n-dimensional to 1-dimensional problems (or words to letters). This makes it possible to estimate the decomposition number in Yao's inequality. For "sum-type" functions this yields often asymptotically optimal results for $C(f_n; 1 \leftrightarrow 2)$. More specifically, lower bounds can be derived with the following inequality.

LEMMA 3. *For any $\mathscr{R} \subset \mathbb{Z}$ and any sum-type function f_n,*

$$C(f_n; 1 \to 2) \geq \log |\mathscr{X}^n||\mathscr{Y}^n| M(f_n)^{-1} \geq \log \left[|\mathscr{X}||\mathscr{Y}| M(f, \mathscr{R}, 1)^{-1} \right]^n.$$

Proof. The first inequality follows from Yao's inequality and (8.4). The second inequality is a consequence of the inequality $M(f_n) \leq M(f, \mathscr{R}, n)$ and Theorem 1.

C. An Upper Bound

In the light of Theorem 4 and the idea for its proof, it is natural to assign to f: $\mathscr{X} \times \mathscr{Y} \to \mathbb{Z}$ the following class $G(f)$ of coarser functions: With every partition $\mathscr{P} = \{S_1 \times T_1, \ldots, S_n \times T_n\}$ of $\mathscr{X} \times \mathscr{Y}$ into rectangles $S_i \times T_i$, $(S_i, T_i) \in \mathscr{P}(f, \{0\}, 1)$, we associate the function $g_{\mathscr{P}}$: $\mathscr{X} \times \mathscr{Y} \to \{1, 2, \ldots, n\}$, where

$$g_{\mathscr{P}}(x, y) = i, \qquad \text{if } (x, y) \in S_i \times T_i. \tag{8.5}$$

$G(f)$ is the set of functions obtained if $\mathscr{P}$ varies over all such partitions.

THEOREM 5. *For a sum-type function f_n,*

$$\overline{\lim_{n \to \infty}} \frac{C(f_n; 1 \leftrightarrow 2)}{n} \leq \min_{g \in G(f)} C(g; 1 \to 2). \tag{8.6}$$

Proof. Let g be such that the minimum is assumed in (8.6) and let $\mathscr{P} = \{S_1 \times T_1, \ldots, S_k \times T_k\}$ be the partition of $\mathscr{X} \times \mathscr{Y}$ on which g was based. We learned from the proof of Theorem 4 that in case both person $P_{\mathscr{X}}$ and person $P_{\mathscr{Y}}$ know $g^n(x^n, y^n) = (g(x_1, y_1), \ldots, g(x_n, y_n))$, only a relatively small additional amount of information is exchanged for both to know $f_n(x^n, y^n)$. More precisely, if $g^n(x^n, y^n) = \gamma^n = (\gamma_1, \ldots, \gamma_n)$, then by the Decomposition Lemma there are functions φ_γ: $S_\gamma \to \mathbb{Z}$, ψ_γ: $T_\gamma \to \mathbb{Z}$ with

$$f(x, y) = \varphi_\gamma(x) + \psi_\gamma(y) \qquad \text{for } (x, y) \in S_\gamma \times T_\gamma \tag{8.7}$$

and

$$f(x^n, y^n) = \sum_{t=1}^{n} \varphi_{\gamma_t}(x_t) + \psi_{\gamma_t}(y_t)$$

$$\text{for } (x^n, y^n) \in \prod_{t=1}^{n} S_{\gamma_t} \times T_{\gamma_t}. \tag{8.8}$$

Furthermore,

$$\left\| \sum_{t=1}^{n} \varphi_{\gamma_t} \right\| \le |\Phi| n, \quad \left\| \sum_{t=1}^{n} \psi_{\gamma_t} \right\| \le |\psi| n. \tag{8.9}$$

Now, by definition, an exchange of $C(g; 1 \leftrightarrow 2)n$ bits suffices for both persons to learn that $g^n(x^n, y^n) = (\gamma_1, \ldots, \gamma_n)$. Using this knowledge $P_{\mathscr{X}}$ can compute $\sum_{t=1}^{n} \varphi_{\gamma_t}(x_t)$ and $P_{\mathscr{Y}}$ can compute $\sum_{t=1}^{n} \psi_{\gamma_t}(y_t)$. By (8.9) an exchange of $\log |\Phi| n + \log |\psi| n$ bits suffices to inform both persons about $f(x^n, y^n)$. However, since $\lim_{n \to \infty} (1/n) \log |\Phi| |\psi| n^2 = 0$, the proof is complete.

D. Examples of Sum-Type Functions

The following conclusion can be drawn from the foregoing results: If we can find a $g \in G(f)$ such that for some γ, $|g^{-1}(\gamma)| = M(f, \{0\}, 1)$ and Yao's bound for $C(g; 1 \leftrightarrow 2)$ is tight (which actually means that $|g^{-1}(\gamma)| = M(f, \{0\}, 1)$ for all γ), then

$$\lim_{n \to \infty} \frac{C(f_n; 1 \leftrightarrow 2)}{n} = \log \frac{|\mathscr{X}| |\mathscr{Y}|}{M(f, \{0\}, 1)}$$

and we know a protocol achieving this value.

We analyse now the four functions $f^{(1)}, \ldots, f^{(4)}$ defined by the tables

$x \backslash y$	0	1	2	3
0	0	2	2	3
1	1	1	4	3
2	0	5	4	5

$f^{(1)}$

$x \backslash y$	0	1	2	3
0	0	0	2	2
1	1	1	2	2
2	1	1	3	3
3	1	1	3	3

$f^{(2)}$

$x \backslash y$	0	1	2	3	4	5
0	−3	−2	−5	3	4	5
1	4	5	2	1	2	3
2	2	3	0	0	1	2
3	−2	−3	0	0	2	4
4	1	0	3	2	4	6
5	−1	−2	1	4	6	8

$f^{(3)}$

$x \backslash y$	0	1	2	3	4	5
0	0	1	0	3	3	3
1	0	2	2	4	4	4
2	0	3	3	5	5	5

$f^{(4)}$

For the first 3 functions the decompositions are as indicated in the tables

By the principle stated above we get

$$\lim_{n\to\infty}\frac{1}{n}C\left(f_n^{(1)};1\leftrightarrow 2\right)=\log_2 3,$$

$$\lim_{n\to\infty}\frac{1}{n}C\left(f_n^{(2)};1\leftrightarrow 2\right)=1,$$

$$\lim_{n\to\infty}\frac{1}{n}C\left(f_n^{(3)};1\leftrightarrow 2\right)=2.$$

For instance, for $f^{(2)}$ the g corresponding to the decomposition drawn is known in values to $P_{\mathcal{Y}}$, who can inform $P_{\mathcal{X}}$ with 1 bit, that is, by sending a 0, if $y\in\{0,1\}$ and a 1, if $y\in\{2,3\}$. Applying the principle literally to $f^{(1)}$ we get the upper bound $\lceil\log_2 3\rceil$.

However, by replacing g by g^m and considering n's divisible by m, as $m\to\infty$ we obtain the bound $\log_2 3$. We leave the analysis of $f^{(3)}$ to the reader. It is more interesting to look at $f^{(4)}$. A straightforward application of our principle gives

$$1=\log\frac{6\cdot 3}{9}\leq\underline{\lim_{n\to\infty}}\frac{C\left(f_n^{(4)};1\leftrightarrow 2\right)}{n}$$

$$\leq\overline{\lim_{n\to\infty}}\frac{C\left(f_n^{(4)};1\leftrightarrow 2\right)}{n}\leq\log 3. \tag{8.10}$$

However, since $f(x,3)=f(x,4)=f(x,5)$ we can identify 3, 4, and 5 and define

x \ y	0 1 2 3
0	0 1 0 3
1	0 2 2 4
2	0 3 3 5

f'

Now $\underline{\lim}_{n\to\infty}(1/n)C(f_n';1\leftrightarrow 2)\geq\log 3$, (8.10) and

$$\lim_{n\to\infty}(1/n)C\left(f_n^{(4)};1\leftrightarrow 2\right)=\lim_{n\to\infty}(1/n)C\left(f_n';1\leftrightarrow 2\right)$$

imply that $\log 3$ is the limiting value.

Finally we comment on the 16 Boolean functions $\{0,1\}\times\{0,1\}\to\{0,1\}$. They fall into 2 groups, namely trivial functions depending on one or

no argument, functions equivalent to symmetric difference Δ, and those equivalent to union $\vee$, which include intersection and difference.

It has been observed by El Gamal and Pang that for d_n: $\{0,1\}^n \times \{0,1\}^n \to \{0,1,\ldots,n\}$ defined by

$$d_n(a^n, b^n) = \sum_{t=1}^{n} a_t \Delta b_t \tag{8.11}$$

$$\lim_{n\to\infty} \frac{1}{n} C(d_n; 1 \leftrightarrow 2) = 1.$$

Obviously, for trivial Boolean functions f_n, $\lim_{n\to\infty}(1/n)C(f_n; 1 \leftrightarrow 2) = 0$. Already (8.3) and (8.4), in conjunction with Corollary 4, give

$$\lim_{n\to\infty} \frac{1}{n} C(\mu_n; 1 \leftrightarrow 2) = 1.$$

That means that for all nontrivial Boolean functions 1-way communication is asymptotically as efficient as 2-way communication. We conjecture this to hold in an exact sense.

References

1. R. Ahlswede, A. El Gamal, and K. F. Pang, A two-family extremal problem in Hamming space, *Discrete Math.* **49** (1984), 1–5.
2. P. Delsarte and P. Piret, An extension of an inequality by Ahlswede, El Gamal, and Pang for pairs of binary codes, *Discrete Math.* **55** (1985), 313–315.
3. J. I. Hall and J. H. van Lint, Constant distance code pairs, *Proc. Kon. Nederl. Akad. Wetensch. Ser. A* **88**, No. 1 (1985), 41–45.
4. R. Ahlswede and M. Moers, Inequalities for code pairs, *European J. Combin.*, in press.
5. A. Yao, Some complexity questions related to distributive computing, *in* "Proceedings 11th Annu. ACM Sympos. Theory of Computing, 1979," pp. 209–219.
6. A. El Gamal and K. F. Pang, Communication complexity of computing the Hamming distance, *SIAM J. Comput.* (1986), 932–947.
7. N. Cai, A bound of sizes of code pairs satisfying the strong 4-words property for Lee distance, *J. System Sci. Math. Sci.* **6**, No. 2 (1986), 129–135.
8. R. Ahlswede, On code pairs with specified Hamming distances, *in* "Proceedings, Conf. on Irregularity of Partitions, Fertöd, Hungary, 1986" (V. Sós, Ed.).
9. H. Abelson, Lower bounds on inform. transfer in distributed computations, *in* "Proceedings, IEEE 19th Annual Sympos. on Foundations of Computer Sci.," Ann Arbor, 1978, pp. 151–158.
10. C. H. Papadimitriou and M. Sipser, Communication complexity, *in* "Proceedings, 14th Annu. ACM Symp. on Theory of Computing, 1982," pp. 201–214.
11. R. Ahlswede, Coloring hypergraphs: A new approach to multiuser source coding I, *J. Combin. Inform. System Sci.* **4**, No. 1 (1979), 76–115; Part II, **5**, No. 3 (1980), 220–268.