Three results on interactive communication

Moni Naor * Alon Orlitsky† Peter Shor†

Abstract

X and Y are random variables. Person $P_{\mathcal{X}}$ knows X, Person $P_{\mathcal{Y}}$ knows Y, and both know the underlying probability distribution of the *random pair* (X, Y). Using a predetermined protocol, they exchange messages over a binary, error-free, channel in order for $P_{\mathcal{Y}}$ to learn X. $P_{\mathcal{X}}$ may or may not learn Y. $\hat{C}_m$ is the number of information bits that must be transmitted (by both persons) in the worst case if only m messages are allowed. $\hat{C}_\infty$ is the corresponding number of bits when there is no restriction on the number of messages exchanged. We consider three aspects of this problem.

$\hat{\mathbf{C}}_\mathbf{4}$. It is known that one-message communication may require exponentially more bits than the minimum possible: for some random pairs, $\hat{C}_1 = 2^{\hat{C}_\infty - 1}$. Yet just two messages suffice to reduce communication to almost the minimum: for all random pairs, $\hat{C}_2 \leq 4\hat{C}_\infty + 3$. We show that, asymptotically, four messages require at most three times the minimum number of bits: for all random pairs, $\hat{C}_4 \leq 3\hat{C}_\infty + o(\hat{C}_\infty)$.

Balanced pairs. Let $\hat{\mu}$ be the maximum number of X values possible with a given Y value, and let $\hat{\eta}$ be the maximum number of Y values possible with a given X value. A random pair is *balanced* if $\hat{\mu} = \hat{\eta}$. It is known that for all balanced pairs, three messages require at most $\log \hat{\mu} + o(\log \hat{\mu})$ bits, hence are asymptotically optimum. It was not known whether two messages are asymptotically optimum. We show that for every c and positive ϵ there is a balanced pair such that $\hat{C}_2 \geq (2 - \epsilon)\hat{C}_\infty \geq c$. Asymptotically, this is the largest possible discrepancy.

Amortized complexity. The amortized complexity of (X, Y) is the limit, as k grows, of the number of bits required in the worst case for k independent repetitions of (X, Y), normalized by k. We show that the four-message amortized complexity of all random pairs is exactly $\log \hat{\mu}$. Hence, when a random pair is repeated many times, no bits can be saved if $P_{\mathcal{X}}$ knows Y in advance.

*Department of Computer Science and Applied Mathematics, Weizmann Institute, Rehovot 76100, Israel (Work done while at IBM Almaden Research Center).

†AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, NJ 07974.

1 Introduction

We provide some background material, describe the three problems considered, and give a few definitions.

1.1 Background

Interactive communication is concerned with various aspects of the following problem. There are two *communicators*: an *informant* $P_{\mathcal{X}}$ having a random variable X and a *recipient* $P_{\mathcal{Y}}$ having a random variable Y. The *random pair* (X,Y) is distributed according to some probability distribution that is known to both communicators. $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ want the recipient, $P_{\mathcal{Y}}$, to learn X with no probability of error. The informant, $P_{\mathcal{X}}$, may or may not learn Y.

To that end $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ alternate in transmitting *messages*: finite sequences of bits. Messages are transmitted over an error-free channel and are determined by an agreed-upon, deterministic, protocol. For every *input* — a possible value assignment for X and Y — the protocol determines a finite sequence of transmitted messages. The protocol is *m-message* if, for all inputs, the number of messages transmitted is at most m.

The *worst-case complexity* of a protocol is the number of bits it requires both communicators to transmit, maximized over all inputs. $\hat{C}_m$, the *m-message complexity*[1] of (X,Y), is the minimum complexity of an m-message protocol for (X,Y). For example, $\hat{C}_1$, the *one-way complexity* of (X,Y), is the number of bits required in the worst-case when $P_{\mathcal{Y}}$ cannot transmit to $P_{\mathcal{X}}$. $\hat{C}_2$ is the number of bits required in the worst-case when at most two messages are permitted: $P_{\mathcal{Y}}$ transmits a message reflecting Y, then $P_{\mathcal{X}}$ responds with a message from which $P_{\mathcal{Y}}$ must infer X. Since empty messages are allowed, $\hat{C}_m$ is a decreasing function of m bounded below by 0. We can therefore define $\hat{C}_\infty$, the *unbounded-message complexity of* (X,Y), to be the limit of $\hat{C}_m$ as $m \to \infty$. It is the *minimum* number of bits that must be transmitted for $P_{\mathcal{Y}}$ to know X, even if no restrictions are placed on the number of messages exchanged. Clearly, for all random pairs,

$$\hat{C}_1 \geq \hat{C}_2 \geq \hat{C}_3 \geq \ldots \geq \hat{C}_\infty \ .$$

Interactive communication can therefore be viewed as a variation of communication complexity [1, 2]. The function computed by $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ is trivial: $f(x,y) = x$, and the difficulty

[1]The pair (X,Y) is implicit in the notation and will be implied by the context.

arises because the inputs x and y are correlated. The following example, taken from [3], illustrates some of the concepts involved.

Example 1 A league has t teams. $P_{\mathcal{Y}}$ knows two teams that played in a game, and $P_{\mathcal{X}}$ knows the team that won the game. They communicate in order for $P_{\mathcal{Y}}$ to learn the winning team.

If only one message is allowed, necessarily from $P_{\mathcal{X}}$ to $P_{\mathcal{Y}}$, it must be based solely on the winner (for that is all $P_{\mathcal{X}}$ knows). If the message transmitted when team i wins is the same as (or a prefix of) the message transmitted when team j wins, then in the event of a match between teams i and j, $P_{\mathcal{Y}}$ cannot tell who the winner is (or when the message ends). Therefore, there must be t different, prefix free, messages and at least one of them must be of length $\geq \lceil \log t \rceil$. This bound is clearly achievable, hence,

$$\hat{C}_1 = \lceil \log t \rceil \ .$$

If two messages are allowed, $P_{\mathcal{Y}}$ considers the binary representations of the two teams that played and transmits $\lceil \log \log t \rceil$ bits describing the location of the first bit where they differ. $P_{\mathcal{X}}$ responds by transmitting a single bit describing the bit value of the winning team in that location. Therefore, $\hat{C}_2 \leq \lceil \log \log t \rceil + 1$. It can be shown that for this example

$$\hat{C}_2 = \ldots = \hat{C}_\infty = \lceil \log \log t \rceil + 1 \ . \qquad \square$$

In this paper we consider three aspects of interactive communication.

1.2 Four messages require at most three times the minimum number of bits

Example 1 shows that for some random pairs, one-message may require exponentially more bits than the minimum necessary:

$$\hat{C}_1 = 2^{\hat{C}_\infty - 1} \ .$$

Yet results in [3] show that two messages always suffice to reduce communication to at most four times the minimum: for all random pairs,

$$\hat{C}_2 \leq 4\hat{C}_\infty + 3 \ .$$

This contrasts with communication complexity where a succession of papers [4, 5, 6] showed that for every m there is a function whose m-message complexity is almost exponentially higher than its $(m+1)$-message complexity. Two messages are *not* always optimum. As shown in [7], for every c and positive ϵ, there are random pairs such that

$$\hat{C}_2 \geq (2-\epsilon)\hat{C}_\infty \geq c \ .$$

One of the main open problems in interactive communication is whether there is an m such that m messages are asymptotically optimum, namely, for all random pairs

$$\hat{C}_m \leq \hat{C}_\infty + o(\hat{C}_\infty) \ .$$

In Section 2 we make a small step towards a resolution of this problem. We show that asymptotically four messages require at most three times the minimum number of bits: for all random pairs

$$\hat{C}_4 \leq 3\hat{C}_\infty + o(\hat{C}_\infty) \ .$$

We remark that for *average-case* complexity four messages are asymptotically optimum [8].

1.3 Balanced pairs: two messages may require twice the minimum number of bits

Let (X,Y) be a random pair. Its *support set* is the set S of possible inputs. The support set is of interest as it determines the m-message complexity $\hat{C}_m$ for all m. $P_\mathcal{Y}$'s *ambiguity* when he has the value y is

$$\mu(y) \stackrel{\text{def}}{=} |\{x : (x,y) \in S\}| \ ,$$

the number of possible X values when $Y=y$. $P_\mathcal{Y}$'s *maximum ambiguity* is

$$\hat{\mu} \stackrel{\text{def}}{=} \max_y\{\mu(y)\}, \tag{1}$$

the maximum number of X values possible with any given Y value. $P_\mathcal{X}$'s ambiguity $\eta(x)$ when he has the value x, and his maximum ambiguity $\hat{\eta}$, are similarly defined. In the league problem of Example 1, $\hat{\mu}=2$ as for every game known to $P_\mathcal{Y}$ there are two possible winners known to $P_\mathcal{X}$. Similarly, $\hat{\eta}$ in that case is $t-1$, corresponding to the number of possible losing teams.

A random pair is *balanced* if $P_\mathcal{X}$ and $P_\mathcal{Y}$ have the same maximum ambiguity:

$$\hat{\mu} = \hat{\eta} \ .$$

Balanced pairs arise naturally whenever X and Y are derived from a single variable or are known to be within some 'distance' from each other. For example:

1. X and Y, inaccurate measurements of the same quantity, are integers within a bounded absolute difference from each other.

2. X and Y, obtained from a noisy binary transmission or from a faulty memory, are n-bit strings within a bounded Hamming distance from each other.

3. X and Y, modified versions of the same file, are binary strings within a small edit distance from each other.

It is shown in [9] that for all balanced random pairs, one-way communication requires at most twice the minimum number of bits:

$$\hat{C}_1 \leq 2\hat{C}_\infty + 1 \ . \tag{2}$$

This bound is almost tight. For all c there is a balanced pair such that

$$\hat{C}_1 \geq 2\hat{C}_\infty - 6 \geq c \ .$$

The most interesting result in [9] is that for all balanced pairs,

$$\hat{C}_3 \leq \log \hat{\mu} + 3 \log\log \hat{\mu} + 11 \ .$$

Hence, although the informant, $P_{\mathcal{X}}$, does not know Y, the number of bits needed to convey X to $P_{\mathcal{Y}}$ is only negligibly larger than would be required if $P_{\mathcal{X}}$ knew Y in advance. Furthermore, three messages are asymptotically optimum: for all balanced pairs,

$$\hat{C}_3 \leq \hat{C}_\infty + 3 \log \hat{C}_\infty + 11.$$

However, it was not known whether *two* messages are asymptotically optimum. In view of the potential practical applications of balanced pairs, it is interesting to determine the minimum number of messages required to convey them efficiently. In Section 3 we show that two messages may require twice the minimum number of bits. For all c and positive ϵ there is a balanced pair such that

$$\hat{C}_2 \geq (2 - \epsilon)\hat{C}_\infty \geq c \ .$$

Considering Inequality (2), this is the largest possible asymptotic discrepancy. Moreover, with this result, we can establish the largest discrepancies between all m-message complexities for balanced pairs.

1.4 Amortized complexity

Feder, Kushilevitz, and Naor [10] considered the following modification of the league problem.

Example 2 Consider two leagues — one for basketball, the other for baseball — each with t teams. $P_{\mathcal{Y}}$ knows two games, one from each league, and $P_{\mathcal{X}}$ knows the two winners of these games. The games and winners are independent of each other. How many bits must be exchanged in the worst case for $P_{\mathcal{Y}}$ to learn both winners?

Example 1 shows that learning the winner of a single game requires exactly $\lceil \log \log t \rceil + 1$ bits in the worst case. Thus, if treated independently, the two games require $2(\lceil \log \log t \rceil + 1)$ bits in the worst case. By considering the two games together [10] showed that there is almost no increase in the number of bits required over the single league case:

$$\hat{C}_2 \leq \lceil \log \log t \rceil + 6 \ .$$

Their proof is repeated when Example 2 is continued in Section 4. □

This modification can be applied to all random pairs. Let (X, Y) be a random pair with support set S. In Section 4 we consider the random pair $(X^{(k)}, Y^{(k)})$ obtained by k independent repetitions of (X, Y). More precisely, $P_{\mathcal{X}}$ knows a sequence $x^{(k)} = (x_1, \ldots, x_k)$ and $P_{\mathcal{Y}}$ knows a sequence $y^{(k)} = (y_1, \ldots, y_k)$ such that $(x_i, y_i) \in S$ for all $i \in \{1, \ldots, k\}$. As before, $P_{\mathcal{Y}}$ wants to learn $x^{(k)}$.

Let $\hat{C}_m^{(k)}$ denote the m-message worst-case complexity of $(X^{(k)}, Y^{(k)})$. We are interested in the behavior of $\hat{C}_m^{(k)}$ for large k. The *m-message worst-case amortized complexity* of (X, Y) is

$$\hat{A}_m \stackrel{\text{def}}{=} \lim_{k \to \infty} \frac{\hat{C}_m^{(k)}}{k} \ ,$$

where the limit exists by subadditivity. Intuitively, $\hat{A}_m$ is the number of bits per repetition of (X, Y) required in the worst case when m messages are allowed. Witsenhausen [11] considered one-message amortized complexity. We are more concerned with the number of bits required when $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ can interact. Using results of Section 2, we find upper bounds on $\hat{C}_4^{(k)}$ and apply them to show that for all random pairs

$$\hat{A}_4 = \ldots = \hat{A}_\infty = \log \hat{\mu}$$

where $\hat{\mu}$ is $P_{\mathcal{Y}}$'s ambiguity defined in Equation (1). Note that $\log \hat{\mu}$ bits are needed even if $P_{\mathcal{X}}$ knows Y in advance. Hence, when a pair is repeated many times, the number of bits per repetition required in the worst case is the same as the number of bits needed if $P_{\mathcal{X}}$ knows Y in advance. Furthermore, this can be achieved using four messages.

1.5 Definitions

In the ensuing sections we will use the following definitions. Let (X, Y) be a random pair with support set S. The *support set* of X is the set

$$S_{\mathcal{X}} \stackrel{\text{def}}{=} \{x : (x, y) \in S \text{ for some } y\}$$

of possible values of X. The *support set* $S_{\mathcal{Y}}$ of Y is similarly defined. Instrumental in analyzing the worst-case complexity measures is G, the *characteristic hypergraph* of (X, Y). Its vertex set is $S_{\mathcal{X}}$ and for every $y \in S_{\mathcal{Y}}$, it contains the hyperedge:

$$E(y) \stackrel{\text{def}}{=} \{x : (x, y) \in S\} \ .$$

The characteristic hypergraph is equivalent to a graph defined by Witsenhausen [11] who considered the one-message version of this problem. For the league problem with t teams, for example, G has t vertices, one corresponding to each (team) value of X. It has $\binom{t}{2}$ edges, one corresponding to each possible (game) value of Y. Each edge contains two vertices (the two possible winning teams in this game). In other words, G is K_t, the complete graph on t vertices.

A coloring of a hypergraph G is an assignment of a color to every vertex of G such that no two vertices sharing a hyperedge are assigned the same color. The chromatic number χ of G is the minimum number of colors required to color G.

2 Four messages require at most three times the minimum number of bits

To show that for all random pairs,

$$\hat{C}_4 \le 3\hat{C}_\infty + o(\hat{C}_\infty) \ , \tag{3}$$

we lower bound $\hat{C}_\infty$ and upper bound $\hat{C}_4$ in terms of $P_{\mathcal{Y}}$'s maximum ambiguity, $\hat{\mu}$, and χ, the chromatic number of the characteristic hypergraph. The lower bound on $\hat{C}_\infty$ is the same one used to prove Corollary 5 in [3]:

Result 1 For all random pairs,

$$\hat{C}_\infty \ge \max\{\log\log\chi, \log\hat{\mu}\} \ . \qquad \square$$

The upper bound on $\hat{C}_4$ is obtained by constructing a low-complexity four-message protocol. To demonstrate the basic design of the protocol we first show how an existing upper bound on $\hat{C}_2$ can be used to obtain a suboptimal upper bound on $\hat{C}_4$.

Result 2 (Theorem 4 in [3]) For all nontrivial[2] random pairs,

$$\hat{C}_2 \leq \log\log\chi + 3\log\hat{\mu} + 4 \ .$$

Outline of proof: $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ can agree in advance on a χ-coloring of G. If $P_{\mathcal{Y}}$ learns the color of X, he can deduce X. We can therefore assume that $S_{\mathcal{X}} = \{1,\ldots,\chi\}$.

A function $f : \{1,\ldots,\chi\} \to \{1,\ldots,\hat{\mu}^2\}$ *perfectly hashes* the ambiguity set $E(y)$ if it is one-to-one over $E(y)$. Perfect-hash functions have been investigated and applied extensively in recent years [12, 13, 14, 15].

It is not hard to show that there is a collection of $\lceil 4\hat{\mu}\log\chi\rceil$ functions such that every $y \in S_{\mathcal{Y}}$ has a function in the collection that perfectly hashes $E(y)$. $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ agree on such a collection and on an encoding of its functions.

$P_{\mathcal{Y}}$, given Y, transmits $\lceil\log\log\chi + \log\hat{\mu} + 2\rceil$ bits identifying a function f in the collection that perfectly hashes $E(y)$. $P_{\mathcal{X}}$ responds with $\lceil 2\log\hat{\mu}\rceil$ bits describing $f(X)$. It is easy to verify that $P_{\mathcal{Y}}$ can deduce X. □

When $\log\log\chi \ll \log\hat{\mu}$, we can reduce the number of transmitted bits by prefixing the protocol with two more messages.

Lemma 1 For all nontrivial random pairs,

$$\hat{C}_4 \leq 2\log\log\chi + 2\log\hat{\mu} + 3\log\log\hat{\mu} + 5 \ .$$

Proof: Let (X,Y) be a random pair. If $\hat{\mu} \leq 31$, the lemma is implied by Result 2. We therefore assume that $\hat{\mu} \geq 32$ and, as in the outline above, that $S_{\mathcal{X}} = \{1,\ldots,\chi\}$.

A function $f : \{1,\ldots,\chi\} \to \{1,\ldots,\hat{\mu}\}$ is *smooth* for $y \in S_{\mathcal{Y}}$ if for all $\alpha \in \{1,\ldots,\hat{\mu}\}$,

$$|f^{-1}(\alpha) \cap E(y)| \leq \log\hat{\mu} \ ,$$

namely, if f assigns any given value to at most $\log\hat{\mu}$ elements of $P_{\mathcal{Y}}$'s ambiguity set $E(y)$. Pick a random function F with the above domain and range. If F is not smooth for $y \in S_{\mathcal{Y}}$, then it assigns the same value to at least $\lfloor\log\hat{\mu}\rfloor + 1$ elements of of $E(y)$. Hence,

$$\Pr(F \text{ is not smooth for } y) < \hat{\mu}\frac{\binom{\hat{\mu}}{\lfloor\log\hat{\mu}\rfloor+1}}{\hat{\mu}^{\lfloor\log\hat{\mu}\rfloor+1}} < \frac{\hat{\mu}}{(\lfloor\log\hat{\mu}\rfloor+1)!} < \frac{1}{4} \ . \tag{4}$$

[2] A random pair is *trivial* if $\hat{\mu} = 1$, namely if X is determined by Y.

Independently pick m such random functions. The probability that none of them is smooth for y is less than $\left(\frac{1}{4}\right)^m$. There are at most $\sum_{i=1}^{\hat{\mu}} \binom{\chi}{i} < \chi^{\hat{\mu}}$ different sets $E(y)$. By the union bound, the probability that there is a $y \in S_{\mathcal{Y}}$ for which no function is smooth is less than $\chi^{\hat{\mu}}(\frac{1}{4})^m$. If $2m \geq \hat{\mu} \log \chi$, then

$$\Pr(\text{there is a } y \text{ for which no function is smooth}) < 1 ,$$

hence there is a collection of $\left\lceil \frac{1}{2}\hat{\mu} \log \chi \right\rceil$ functions that contains a smooth function for every $y \in S_{\mathcal{Y}}$.

$P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ agree on such a collection and on a $\lceil \log\log \chi + \log \hat{\mu} - 1 \rceil$-bit encoding of the functions in the collection. Given Y, $P_{\mathcal{Y}}$ transmits the encoding of a function f that is smooth for y. Then, $P_{\mathcal{X}}$ responds with $\lceil \log \hat{\mu} \rceil$ bits describing the value of $f(X)$.

Now, $P_{\mathcal{X}}$ and $P_{\mathcal{Y}}$ can restrict themselves to a random pair with maximum ambiguity of at most $\log \hat{\mu}$ and chromatic number of at most χ. They can use the two-message protocol of Result 2 above to communicate X to $P_{\mathcal{Y}}$. $\square$

The protocol described by the lemma requires $P_{\mathcal{Y}}$ to transmit $\log\log \chi + \log \hat{\mu}$ bits describing a function f that is smooth for $E(y)$ and then $\log\log \chi + \log\log \hat{\mu}$ bits identifying a function that perfectly hashes $E(y) \cap f^{-1}(\alpha)$. The next theorem shows that some of these bits can be saved.

Let

$$\sigma \stackrel{\text{def}}{=} |\{E(y) : y \in S_{\mathcal{Y}}\}| \tag{5}$$

denote the number of different ambiguity sets, or, equivalently, the number of different edges in the characteristic hypergraph.

Theorem 1 For all nontrivial random pairs,

$$\hat{C}_4 \leq \log\log \sigma + \log \hat{\mu} + 3 \log\log \hat{\mu} + 7 .$$

Proof: Let (X, Y) be a random pair and let I, J, and b be integers. Consider an $(I+1)$ by J array of functions, each defined over $S_{\mathcal{X}}$. The top row consists of functions $f_1, \ldots, f_J$, each with range $\{1, \ldots, \hat{\mu}\}$. The other I rows consist of functions $g_{1,1}, \ldots, g_{I,J}$, each with range $\{1, \ldots, b\}$.

Let $y \in S_{\mathcal{Y}}$. As in the previous lemma, the function f_j is *smooth* for y if for all $\alpha \in \{1, \ldots, \hat{\mu}\}$,

$$|f_j^{-1}(\alpha) \cap E(y)| \leq \log \hat{\mu} .$$

The jth column of the array is *good* for y if

1. f_j is smooth for y

2. For all $\alpha \in \{1,\ldots,\hat{\mu}\}$, there is $i \in \{1,\ldots,I\}$ such that $g_{i,j}$ perfectly hashes $f_j^{-1}(\alpha) \cap E(y)$.

The array is *good* for (X,Y) if every $y \in S_{\mathcal{Y}}$ has a good column. A good array can be used to construct a four-message protocol for (X,Y) as follows:

1. $P_{\mathcal{Y}}$ transmits $\lceil \log J \rceil$ bits describing the index j of a good column for y.

2. $P_{\mathcal{X}}$ transmits $\lceil \log \hat{\mu} \rceil$ bits describing $\alpha \stackrel{\text{def}}{=} f_j(x)$.

3. $P_{\mathcal{Y}}$ transmits $\lceil \log I \rceil$ bits describing an index i such that $g_{i,j}$ perfectly hashes $f_j^{-1}(\alpha) \cap E(y)$.

4. $P_{\mathcal{X}}$ transmits $\lceil \log b \rceil$ bits describing $g_{i,j}(x)$.

It is easy to see that $P_{\mathcal{Y}}$ can use these transmissions to deduce x. The total number of bits transmitted is $\lceil \log J \rceil + \lceil \log \hat{\mu} \rceil + \lceil \log I \rceil + \lceil \log b \rceil$. To prove the theorem, we show that every (X,Y) pair has a good array with $J = \lceil \log \sigma \rceil$, $I = \left\lceil \frac{\log \hat{\mu} + 2}{\log(4/3)} \right\rceil$, and $b = \lceil \log \hat{\mu} \rceil^2$.

Pick all functions at random. We prove that with positive probability the resulting array is good for (X,Y). Column j is "bad" for y if either of the following holds:
(A) f_j is not smooth for y,
(B) f_j is smooth for i but for some $\alpha \in \{1,\ldots,\hat{\mu}\}$, no $g_{i,j}$ perfectly hashes $f_j^{-1}(\alpha) \cap E(y)$.

As with Lemma 1, the theorem is implied by Result 2 for $\hat{\mu} \leq 31$, hence we assume $\hat{\mu} \geq 32$. Inequality (4) showed that

$$\Pr(A) < \frac{1}{4} .$$

If f_j is smooth for y then, for every $\alpha \in \{1,\ldots,\hat{\mu}\}$ and every $i \in \{1,\ldots,I\}$, the probability that $g_{i,j}$ perfectly hashes $E(y) \cap f_j^{-1}(\alpha)$ is

$$\frac{b^{\underline{\log \mu(y)}}}{b^{\log \mu(y)}} \geq \frac{b^{\underline{\log \hat{\mu}}}}{b^{\log \hat{\mu}}} > \left(\frac{b - \log \hat{\mu}}{b}\right)^{\log \hat{\mu}} \geq \left(1 - \frac{1}{\log \hat{\mu}}\right)^{\log \hat{\mu}} \geq \frac{1}{4} ,$$

where $b^{\underline{k}} \stackrel{\text{def}}{=} b \cdot (b-1) \cdot \ldots \cdot (b-k+1)$ denotes the kth falling power of b. Hence,

$$\Pr(B) < \hat{\mu}\left(\frac{3}{4}\right)^I \leq \frac{1}{4} .$$

Therefore,

$$\Pr(\text{column } j \text{ is bad for } y) < \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

The array has J columns, hence the probability that none is good for y is at most $\left(\frac{1}{2}\right)^J$. There are σ different sets $E(y)$, hence,

$$\Pr(\text{there is a } y \text{ with no good column}) < \left(\frac{1}{2}\right)^J \sigma \leq 1 \ ,$$

so there must be a good array. □

Corollary 1 For all nontrivial random pairs,

$$\hat{C}_4 \leq \log\log\chi + 2\log\hat{\mu} + 3\log\log\hat{\mu} + 7 \ .$$

Proof: Again, we can assume that $S_{\mathcal{X}} = \{1,\ldots,\chi\}$, hence the number σ of different sets $E(y)$ is at most $\sum_{i=1}^{\hat{\mu}} \binom{\chi}{i} < \chi^{\hat{\mu}}$. □

Combining the corollary with Result 1, we obtain

Corollary 2 For all random pairs,

$$\hat{C}_4 \leq 3\hat{C}_\infty + o(\hat{C}_\infty) \ .$$ □

3 Balanced distributions: two messages may require twice the minimum number of bits

Let $G = (V, E)$ be a hypergraph. The hypergraphs $G_1 = (V, E_1), \ldots, G_k = (V, E_k)$ *edge cover* G if every edge in G belongs to at least one G_i. Namely,

$$E \subseteq \bigcup_{i=1}^{k} E_i \ .$$

The *chromatic-decomposition number* of G is

$$\mathcal{D}(G) \stackrel{\text{def}}{=} \min\left\{\sum_{i=1}^{k} \chi(G_i) : G_1, \ldots, G_k \text{ edge cover } G\right\},$$

the minimum sum of the individual chromatic numbers in any edge cover of G. The following result relates the chromatic-decomposition number of a random pair and its two-message complexity.

Result 3 (Lemma 2 in [7]) If for all $y_1, y_2 \in S_{\mathcal{Y}}$ there exists an $x \in S_{\mathcal{X}}$ such that $(x, y_1), (x, y_2) \in S$ then

$$\hat{C}_2 \geq \lceil \log \mathcal{D}(G) \rceil \ .$$ □

We use this result to show that a randomly chosen balanced (X,Y) pair will almost surely have:

$$\hat{C}_2 \geq 2\hat{C}_\infty - o(\hat{C}_\infty) \ .$$

Asymptotically, this discrepancy is the largest possible. Equation (2) implies that for all random pairs, $\hat{C}_2 \leq \hat{C}_1 \leq 2\hat{C}_\infty + 1$.

Let $S_\mathcal{X} = S_\mathcal{Y} = \{1,\ldots,n^2\}$. Generate the support set S at random by admitting each input $(x,y) \in \{1,\ldots,n^2\} \times \{1,\ldots,n^2\}$ independently with probability $\frac{1}{n}$. A 'typical' support set will therefore have roughly n possible y's for every value of x and roughly n possible x's for every value of y. Correspondingly, the characteristic hypergraph G has n^2 vertices and n^2 edges. Each edge contains any given vertex with probability $\frac{1}{n}$. 'Typically,' the graph has roughly n vertices in every edge and roughly n edges incident on every vertex.

Lemma 2 With probability approaching 1 as n increases (*w.p.a.1*), the random characteristic hypergraph G generated above satisfies:

$$\mathcal{D}(G) \geq \frac{n^2}{16 \ln n} \ .$$

Proof: We show that w.p.a.1 , for all $8n\ln n \leq m \leq n^2$, all sets of m edges have chromatic number of at least $\frac{m}{8\ln n}$. Since w.p.a.1 all edges have at least $n/2$ vertices, we obtain that w.p.a.1 for *all* m, all subsets of m edges have chromatic number of at least $\frac{m}{16\ln n}$. The theorem then follows because if $G_1,\ldots,G_k$ is an edge cover of G, where G_i has m_i edges, then w.p.a.1

$$\sum_{i=1}^{k} \chi(G_i) \geq \sum_{i=1}^{k} \frac{m_i}{16\ln n} \geq \frac{n^2}{16\ln n} \ .$$

Let m be an integer satisfying $8n\ln n \leq m \leq n^2$. and let x be an integer satisfying $n \leq x \leq \frac{m}{8\ln n}$. We show that w.p.a.1 every set of m edges is not x colorable.

Fix an x-coloring of the vertices, namely a mapping from $\{1,\ldots,n^2\}$ to $\{1,\ldots,x\}$. For $i \in \{1,\ldots,x\}$ let α_i be the number of vertices colored i. For any color $i \in \{1,\ldots,x\}$,

$$\begin{aligned}
\Pr(\text{a given edge contains at most one vertex colored } i) &= \left(1-\frac{1}{n}\right)^{\alpha_i} + \frac{\alpha_i}{n}\left(1-\frac{1}{n}\right)^{\alpha_i - 1} \\
&= \left(1-\frac{1}{n}\right)^{\alpha_i}\left(1+\frac{\alpha_i}{n-1}\right) .
\end{aligned}$$

Disjoint vertex sets are included in an edge independently, hence

$$\Pr(\text{a given edge is well colored}) = \prod_{i=1}^{x}\left(\left(1-\frac{1}{n}\right)^{\alpha_i}\left(1+\frac{\alpha_i}{n-1}\right)\right)$$

$$
\begin{aligned}
&= \left(1-\frac{1}{n}\right)^{n^2} \prod_{i=1}^{x}\left(1+\frac{\alpha_i}{n-1}\right) \\
&\leq \left(1-\frac{1}{n}\right)^{n^2} \left(1+\frac{n^2}{x(n-1)}\right)^{x} \\
&\leq \left(e^{-\frac{1}{n}-\frac{1}{2n^2}}\right)^{n^2} \left(e^{\frac{n^2}{x(n-1)}-\frac{n^4}{2x^2(n-1)^2}+\frac{n^6}{3x^3(n-1)^3}}\right)^{x} \\
&= e^{-\frac{n^4}{2x(n-1)^2}+\frac{n^6}{3x^2(n-1)^3}+\frac{n}{n-1}-\frac{1}{2}} \\
&\leq e^{-\frac{n^2}{4x}}
\end{aligned}
$$

where the first inequality follows from convexity arguments, and the last — from our assumption that $n \leq x \leq \frac{m}{8\ln n} \leq \frac{n^2}{8\ln n}$ (and that n is sufficiently large).

As each edge is chosen independently,

$$
\Pr(\text{a given set of } m \text{ edges is well colored}) \leq e^{-\frac{n^2 m}{4x}} \; .
$$

There are x^{n^2} x-colorings of G and $\binom{n^2}{m} \leq 2^{n^2}$ sets of m edges. Hence,

$$
\begin{aligned}
\Pr(\text{there are } m \text{ edges that are } x\text{-colorable}) &\leq e^{-n^2(\frac{m}{4x}-\ln(2x))} \\
&\leq e^{-n^2(2\ln n-\ln(n^2/4\ln n))} \\
&= e^{-n^2\ln(4\ln n)} \; .
\end{aligned}
$$

This holds for all $n^2 - 8n\ln n$ values of m in the required range, hence the probability that there is an $m \geq 8n\ln n$ and a set of m edges that is x-colorable goes to zero. □

Theorem 2 For every $c > 0$, there is a balanced (X, Y) pair such that

$$
\hat{C}_2 \geq 2\hat{C}_3 - o(\hat{C}_3) \geq c \; .
$$

Proof: Essentially, the result is already proven. Generate an (X, Y) pair as described prior to Lemma 2. With probability approaching 1, $\hat{\mu}$ and $\hat{\eta}$ are about n, while $\mathcal{D}(G) \geq \frac{n^2}{16\ln n}$. Hence, by Theorem 1 in [16],

$$
\hat{C}_3 \leq \log n + 3\log\log n + 11 \; , \tag{6}
$$

whereas Result 3 implies

$$
\hat{C}_2 \geq \log \mathcal{D}(G) \geq 2\log n - \log\ln n - 4 \; . \tag{7}
$$

There are however two technical difficulties. First, the conditions of Result 3 may not be met. G may have two edges that do not intersect. To overcome this, we can increase the probability of admitting each point to S from $\frac{1}{n}$ to $\frac{\log n}{n}$. Now the probability that there are two non intersecting edges is at most

$$\binom{n^2}{2}\left(1-\left(\frac{\log n}{n}\right)^2\right)^{n^2} \leq n^{4-\log n}$$

Hence, w.p.a.1 the new pair satisfies the conditions of Result 3 and has the properties described by Lemma 2

The second technical problem is that the (new as well as original) pair may not be balanced: since it is picked randomly, $\hat{\mu}$ need not exactly equal $\hat{\eta}$. This can be easily fixed. With probability approaching 1, the new pair has

$$\max\{\hat{\mu}, \hat{\eta}\} \leq 2n \log n \ .$$

We can augment its support set by adding inputs in a single row or a single column to equalize $\hat{\mu}$ and $\hat{\eta}$. This will only marginally effect the upper bound $\hat{C}_3$, and can only increase $\hat{C}_2$. Hence, updating Equations (6) and (7), we see that w.p.a.1 the resulting random pair satisfies:

$$\hat{C}_2 \geq 2\hat{C}_3 - 9 \log \hat{C}_3 - 30 \ . \qquad \square$$

With this result, we can completely characterize the largest discrepancy between the m-message complexities of balanced random pairs. For $i < j$, let

$$R_{i,j} \stackrel{\text{def}}{=} \lim_{\hat{C}_j \to \infty} \sup \left\{ \frac{\hat{C}_i}{\hat{C}_j} : (X, Y) \text{ is balanced} \right\}$$

be the largest asymptotic ratio between $\hat{C}_i$ and $\hat{C}_j$ for balanced (X, Y) pairs. Then,

$$R_{i,j} = \begin{cases} 2 & \text{if } i = 1 \text{ or } 2 \\ 1 & \text{otherwise.} \end{cases}$$

Namely,

1. The largest asymptotic ratio between $\hat{C}_1$ and $\hat{C}_m$ is 2 for all $m \geq 2$. This ratio is achieved for all $m \geq 2$ by the projective plane problem (Lemma 2 in [16]), and for $m \geq 3$ by the rooks problem (Example 2 in [16], see also open problems in Section 5).

2. As we just proved, the largest asymptotic ratio between $\hat{C}_2$ and $\hat{C}_m$ is 2 for all $m \geq 3$.

3. Since three messages are asymptotically optimal for all balanced pairs, the largest asymptotic ratio between $\hat{C}_3$ and $\hat{C}_m$ is 1 for all $m \geq 4$.

4 Amortized complexity

We begin with Feder, Kushilevitz, and Naor's [10] proof that the two-leagues problem requires at most five more bits than a single-league problem.

Example 2 (continued) Consider a binary array with t rows and J columns. View the ith row in the array as a mapping assigning a J-bit sequence to the ith team (in both leagues[3]). The jth column of the array is *good* for basketball game (i, i') and baseball game (k, k') if for both games, the two teams have different bits in that column. Namely, if $a_{i,j} \neq a_{i',j}$ and $a_{k,j} \neq a_{k',j}$. The array is *good* if every pair of games has a good column. A good array can be used to construct a two-message protocol requiring $\lceil \log J \rceil + 2$ bits:

1. $P_{\mathcal{Y}}$ transmits $\lceil \log J \rceil$ bits describing a column that is good for both games.

2. $P_{\mathcal{X}}$ responds with the bit that the winning basketball team has in that column followed by the bit that the winning baseball team has in the column.

We show that there is a good array with $J = \left\lceil \frac{4 \log t}{\log(4/3)} \right\rceil$, hence,

$$\hat{C}_2 \leq \left\lceil \log \frac{4 \log t}{\log(4/3)} \right\rceil + 2 \leq \lceil \log \log t \rceil + 6 \ .$$

Pick the array at random.

$$\Pr(\text{a given column is bad for a given pair of games}) \leq \frac{3}{4}.$$

Since columns are chosen independently,

$$\Pr(\text{all columns are bad for a given pair of games}) \leq \left(\frac{3}{4}\right)^J \ .$$

There are $\binom{t}{2}^2$ pairs of games, hence

$$\Pr(\text{there is a pair of games for which all columns are bad}) \leq \left(\frac{3}{4}\right)^J \binom{t}{2}^2 < 1 \ .$$

Hence there must be a good array. $\square$

We now determine the amortized complexity of an arbitrary random (X, Y) pair. Let $(X^{(k)}, Y^{(k)})$ be the random pair obtained by k independent repetitions of (X, Y). Formally, $x^{(k)} = (x_1, \ldots, x_k)$, $y^{(k)} = (y_1, \ldots, y_k)$, and

$$S^{(k)} \stackrel{\text{def}}{=} \{(x^{(k)}, y^{(k)}) : (x_i, y_i) \in S \text{ for all } i \in \{1, \ldots, k\}\}.$$

[3]In general, one can assign different arrays with the different leagues, but this is unnecessary here.

Let $G^{(k)}$ be the characteristic hypergraph of $(X^{(k)}, Y^{(k)})$. It is easy to verify that $G^{(k)}$ is the *k-fold strong direct product* of G with itself. Its vertex set is the k-fold cartesian product $S_{\mathcal{X}} \times \cdots \times S_{\mathcal{X}}$, and for every k-tuple $(e_1, \ldots, e_k)$ of G-edges, $G^{(k)}$ contains the cartesian-product edge $e_1 \times \cdots \times e_k$. Let $\hat{\mu}^{(k)}$ and $\sigma^{(k)}$ be $P_{\mathcal{Y}}$'s maximum ambiguity and the number of different ambiguity sets of $(X^{(k)}, Y^{(k)})$. It follows that:

1. $\hat{\mu}^{(k)} = \hat{\mu}^k$
2. $\sigma^{(k)} = \sigma^k$.

Therefore, Theorem 1 implies:

Corollary 3 For all nontrivial random pairs and all k

$$\hat{C}_4^{(k)} \leq k \log \hat{\mu} + 4 \log k + \log\log \sigma + 3 \log\log \hat{\mu} + 7 \ . \qquad \square$$

This can be readily used to determine the amortized complexity:

Corollary 4 For all nontrivial random pairs,

$$\hat{A}_\infty = \ldots = \hat{A}_4 = \log \hat{\mu} \ .$$

Proof: Corollary 3 provides the upper bound and Result 1 yields the lower bound:

$$\hat{C}_\infty^{(k)} \geq \log \hat{\mu}^{(k)} = k \log \hat{\mu} \ . \qquad \square$$

5 Open problems

5.1 m-message complexity

One of the main unsolved question concerning interactive communication remains: Is there an m such that m messages are asymptotically optimum, namely, for all random pairs

$$\hat{C}_m \leq \hat{C}_\infty + o(\hat{C}_\infty) \ . \tag{8}$$

Note that for *average-case* complexity four messages are asymptotically optimum. Let $\bar{C}_m$ be the m-message average-case complexity. Then for all random pairs [8],

$$\bar{C}_4 \leq \bar{C}_\infty + 3 \log \bar{C}_\infty + 13.5 \ .$$

It seems that an essential element in proving (8) is a lower bound on $\hat{C}_\infty$ that improves on Result 1 (Section 2).

5.2 Balanced pairs

We have proved that there exists a balanced pair for which two messages require twice the minimum number of bits. It would be interesting to demonstrate a specific balanced pair satisfying $\hat{C}_2 \geq (1+\epsilon)\hat{C}_3$ for some positive ϵ. Following are two pairs with an unknown two-message complexity, hence with a potential for such a discrepancy.

Example 3 (Rooks) Consider an n by n chessboard with two rooks in mutually capturing positions (i.e., in the same row, or the same column, or both (same position)). $P_{\mathcal{X}}$ knows the position X of one rook. $P_{\mathcal{Y}}$ knows the position Y of the other rook and wants to learn X. Formally,

$$S \stackrel{\text{def}}{=} \{((x_1,x_2),(y_1,y_2)) : 0 \leq x_1, x_2, y_1, y_2 \leq n-1 \text{ and either } x_1 = y_1 \text{ or } x_2 = y_2\} \ .$$

It was shown in [16] that

$$\hat{C}_1 = \lceil 2 \log n \rceil$$

and that

$$\log n \leq \hat{C}_\infty \leq \cdots \leq \hat{C}_3 \leq \log n + \log\log n + 1 \ .$$

Little is known about $\hat{C}_2$. A simple two-message protocol, whose description we omit, shows that

$$\hat{C}_2 \leq 1.5 \log n + 4 \ .$$

Can better bounds be derived? □

Example 4 (Unproportional intersection) X and Y are disjoint $\lfloor \frac{n}{3} \rfloor$-element subsets of $\{1,\ldots,n\}$.

For every two X-sets, there is a Y-set that is disjoint from both, hence

$$\hat{C}_1 = \log \binom{n}{\lfloor n/3 \rfloor} \approx .9183n \ .$$

Yet (X,Y) is balanced, hence

$$\hat{C}_3 \approx \log \hat{\mu} = \log \binom{\lceil 2n/3 \rceil}{\lfloor n/3 \rfloor} \approx \frac{2}{3}n \ ,$$

with a similar number of bits needed even if $P_{\mathcal{X}}$ knows Y in advance. What is $\hat{C}_2$?

More generally, let $k \leq m \leq n$ be integers. X and Y are two m-element subsets of $\{1,\ldots,n\}$ such that $|X \cap Y| = k$. The pair (X,Y) is balanced, hence $\hat{C}_3$ is roughly the number of bits needed if $P_{\mathcal{X}}$ knew Y in advance: $\log \binom{k}{l} + \log \binom{n-k}{k-l}$. What is $\hat{C}_2$? □

5.3 Amortized complexity

We have shown that $\hat{A}_4 = \ldots = \hat{A}_\infty = \log \hat{\mu}$. Can $\hat{A}_m$ be determined for $m \leq 3$?

For one-way amortized complexity, let $\chi^{(k)}$ be the chromatic number of $G^{(k)}$, then

$$\hat{A}_1 = \lim_{k \to \infty} \frac{1}{k} \log \chi^{(k)} .$$

Hence results on the chromatic number of strong direct products of graphs translate directly to results on amortized complexity.

Witsenhausen [11] showed that when the characteristic hypergraph is the pentagon graph, $\hat{C}_1 = 2$ while $\hat{A}_1 = \frac{1}{2} \log 5$. Feder, Kushilevitz, and Naor [10] applied results of Linial and Vazirani [17] on the chromatic numbers of strong direct products of graphs to obtain bounds on a related problem.

What about $\hat{A}_2$ and $\hat{A}_3$? Result 2 can be refined to include σ, the number of $P_\mathcal{Y}$'s ambiguity sets defined in Equation (5):

$$\hat{C}_2 \leq \log\log\sigma + 2\log\hat{\mu} + 4 .$$

When applied to $(X^{(k)}, Y^{(k)})$, this shows that

$$\hat{A}_2 \leq 2 \log \hat{\mu} .$$

Can this bound be improved?

References

[1] H. Abelson. Lower bounds on information transfer in distributed computations. In *Proceedings of the 19th Annual Symposium on Foundations of Computer Science*, 1978.

[2] A.C. Yao. Some complexity questions related to distributive computing. In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing*, pages 209–213, 1979.

[3] A. Orlitsky. Worst-case interactive communication I: Two messages are almost optimal. *IEEE Transactions on Information Theory*, 36(5):1111–1126, September 1990.

[4] C.H. Papadimitriou and M. Sipser. Communication complexity. In *Proceedings of the 14th Annual ACM Symposium on Theory of Computing*, pages 196–200, 1982.

[5] P. Duris, Z. Galil, and G. Schnitger. Lower bounds on communication complexity. In *Proceedings of the 16th Annual ACM Symposium on Theory of Computing*, pages 81–91, 1984.

[6] N. Nisan and A. Wigderson. Rounds in communication complexity revisited. In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing*, 1991.

[7] A. Orlitsky. Worst-case interactive communication II: Two messages are not optimal. *IEEE Transactions on Information Theory*, 37(4):995–1005, July 1991.

[8] A. Orlitsky. Average-case interactive communication. *IEEE Transactions on Information Theory*, 38(4):1534–1547, July 1992.

[9] A. Orlitsky. Interactive communication of balanced distributions and of correlated files. *SIAM Journal of Discrete Mathematics*, 6(4):548–564, November 1993.

[10] T. Feder, E. Kushilevitz, and M. Naor. Amortized communication complexity. In *Proceedings of the 32nd Annual Symposium on Foundations of Computer Science*, pages 239–248, 1991.

[11] H. Witsenhausen. The zero-error side information problem and chromatic numbers. *IEEE Transactions on Information Theory*, 22(5):592–593, September 1976.

[12] M.L. Fredman, J. Komlós, and E. Szemeredi. Storing a sparse table with $o(1)$ worst case access time. *Journal of the Association for Computing Machinery*, 31:538–544, 1984.

[13] M.L. Fredman and J. Komlós. On the size of separating systems and families of perfect hash functions. *SIAM Journal on Algebraic and Discrete Methods*, 5(1):61–68, 1984.

[14] J. Friedman. Constructing $o(n \log n)$ size monotone formula for the kth threshold function on n variables. *SIAM J Computing*, 15:641–654, 1986.

[15] J. Körner. Fredman-Komlós bounds and information theory. *SIAM. J. Alg. Disc. Meth.*, 7:560–570, 1986.

[16] A. Orlitsky. Interactive communication: Balanced distributions, correlated files, and average-case complexity. In *Proceedings of the 32nd Annual Symposium on Foundations of Computer Science*, pages 228–238, 1991.

[17] N. Linial and U. Vazirani. Graph products and chromatic numbers. In *Proceedings of the 30th Annual Symposium on Foundations of Computer Science*, pages 124–128, 1989.