Average-case interactive communication

Alon Orlitsky
AT&T Bell Laboratories

Abstract

$X$ and $Y$ are random variables. Person $P_X$ knows $X$, Person $P_Y$ knows $Y$, and both know the joint probability distribution of the pair $(X, Y)$. Using a predetermined protocol, they communicate over a binary, error-free, channel in order for $P_Y$ to learn $X$. $P_X$ may or may not learn $Y$. How many information bits must be transmitted (by both persons) on the average?

At least $H(X|Y)$ bits must be transmitted and $H(X) + 1$ bits always suffice\(^1\). If the support set of $(X, Y)$ is a cartesian product of two sets, then $H(X)$ bits must be transmitted. If the random pair $(X, Y)$ is uniformly distributed over its support set, then $H(X|Y) + 3\log(H(X|Y) + 1) + 15.5$ bits suffice. Furthermore, this number of bits is achieved when $P_X$ and $P_Y$ exchange four messages (sequences of binary bits).

The last two results show that when the arithmetic average number of bits is considered: (1) there is no asymptotic advantage to $P_X$ knowing $Y$ in advance; (2) four messages are asymptotically optimum. By contrast, for the worst-case number of bits: (1) communication can be significantly reduced if $P_X$ knows $Y$ in advance; (2) it is not known whether a constant number of messages is asymptotically optimum.

1 Introduction

In the following subsections we describe the problem, its background, and the results derived.

\(^1\) $H(X)$ is the entropy of $X$ and $H(X|Y)$ is the conditional entropy of $X$ given $Y$. Entropies are binary.
1.1 The problem

Consider two communicators: an informant \( P_x \) having a random variable \( X \) and a recipient \( P_y \) having a, possibly dependent, random variable \( Y \). Both communicators want the recipient, \( P_y \), to learn \( X \) with no probability of error, whereas the informant, \( P_x \), may or may not learn \( Y \). To that end they communicate over an error-free channel. How many information bits must be transmitted on the average?

This problem is a variation on a scenario considered by El Gamal and Orlitsky [1] where both communicators, not just \( P_y \), want to learn the other’s random variable. As elaborated on below, some of our results either follow from or extend results therein.

We assume that the communicators alternate in transmitting messages: finite sequences of bits. Messages are determined by an agreed-upon, deterministic, protocol. A formal definition of protocols for the current model is given in [2]. Essentially, a protocol for \((X, Y)\) (i.e., a protocol for transmitting \( X \) to a person who knows \( Y \)) guarantees that the following properties hold. (1) Separate transmissions: each message is based on the random variable known to its transmitter and on previous messages. (2) Implicit termination: when one communicator transmits a message, the other knows when it ends, and when the last message ends, both communicators know that communication has ended. (3) Correct decision: when communication ends, the recipient, \( P_y \), knows \( X \).

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 average complexity of the protocol is the expected number of bits it requires both communicators to transmit (expectation is taken over all inputs). \( C_m(X|Y) \), the \( m \)-message average complexity of \((X, Y)\), is the minimum average complexity of an \( m \)-message protocol for \((X, Y)\). It is the minimum average number of bits transmitted by both communicators using a protocol that never exchanges more than \( m \) messages. Since empty messages are allowed, \( C_m(X|Y) \) is a decreasing function of \( m \) bounded below by 0. We can therefore define \( \bar{C}_\infty(X|Y) \), the unbounded-message complexity of \((X, Y)\), to be the limit of \( C_m(X|Y) \) as \( m \to \infty \). It is the minimum number of bits that must be transmitted on the average for \( P_y \) to know \( X \), even if no restrictions are placed on the number of messages exchanged. In summary,

\[
C_1(X|Y) \geq C_2(X|Y) \geq C_3(X|Y) \geq \cdots \geq \bar{C}_\infty(X|Y) .
\]