Leo Szilard and Unique Decipherability

Consider the classical condition of unique decipherability,

$$\sum 2^{-C(g)} \leq 1, \tag{1}$$

where $C(g)$ is the number of binary letters required to code the word $L(g)$ in a coding procedure.

In the case of prefix codes, i.e., when no word code is also the beginning of another word code, it was reportedly first proved in Kraft's unpublished dissertation that inequality (1) implies the existence of a prefix code. The first printed proof seems to be given in my early paper [1], where I introduced for the first time the concept of error-limitation (also sometimes referred to as "self-synchronization").

Actually, the prefix condition is not necessary for unique decipherability, as was shown by Sardinas and Patterson in 1953. Indeed, my 1954 proof of the inequality (1) showed that it holds whenever there is a uniquely decipherable code. This paper [1] is entirely based upon the properties of the roots of the generating functions associated with the set of costs $C(g)$. Two years later, a similar but more detailed proof was independently presented by McMillan [2]; it is reproduced in several textbooks. Other proofs are also known.

While proving the generalized form of (1), I proposed to call it the Szilard's inequality, because "it was introduced by Szilard (in a structurally identical problem relating to Maxwell's Demons)." A few readers, who must be commended for having looked up Szilard [3] complained that they saw no identity, in fact no connection, between the problems of Maxwell's Demon and those of unique decipherability. Shortly after Szilard's passing, his paper was finally translated (and will hopefully be read as well as referred to). Therefore, it seems appropriate to make my early statement more explicit: such is the purpose of this correspondence.

First Point: The condition (1) is clearly both *necessary and sufficient* in order that

$$\sum p(g)C(g) \geq \sum [-p(g) \log_2 p(g)] \tag{2}$$

for every set of numbers $p(g)$ for which $\sum p(g) = 1$.

Second Point: When the various terms are properly interpreted, (2) may express either Shannon's theorem on noiseless coding, or the second principle of thermodynamics (considered as an inequality between "irreversible" and "reversible" changes of entropy). In this sense, the *necessity* of condition (1) is indeed "structurally identical" in the two contexts as was claimed by Mandelbrot [1].

The *sufficiency* of that condition is a different matter, and the proof requires more assumptions about the concrete context: the condition of decipherability requires a simple argument about coding trees, and the thermodynamical inequality requires one of those marvelous think-experiments which adorn thermodynamics (incidentally, Szilard's proof actually requires not only the second principle of thermodynamics, but also the properties of perfect gases).

Thus, the inequality (1) *does* continue to raise the well-known uncertainty implied in attributions of results to individuals. I continue to favor attributing (1) to Szilard.

BENOIT MANDELBROT
IBM Research Corp.
Yorktown Heights, N. Y.
and
Mass. Inst. Tech.
Cambridge, Mass.

Manuscript received February 1, 1965.

REFERENCES

[1] Mandelbrot, B., On recurrent noise-limiting coding, *Proc. Sym. on Information Networks*, Polytechnic Institute of Brooklyn, N. Y., 1954, pp 205–221.
[2] McMillan, B., Two inequalities implied by unique decipherability, *IRE Trans. on Information Theory*, vol IT-2, Dec 1956, pp 115–116.
[3] Szilard, L., Uber die Entropieverminderung in einem Thermodynamischen System bei Einsgriff intelligenter Wesen, *Z. Phy. (Germany)*, vol 53, 1929, pp 840–856.
[4] Szilard, L., On the decrease of entropy in a thermodynamic system by the intervention of intelligent beings, *Behavioral Science*, vol 9, Oct 1964, pp 301–310.

Reprinted from the IEEE Transactions on Information Theory, Volume IT-11 (July 1965), pages 455 and 456.