166. *On "Amount of Information"*

By Kōmei SUZUKI

(Comm. by K. KUNUGI, M.J.A., Dec. 13, 1956)

We have pursued the mathematical characters of the "amount of information" propounded by Norbert Wiener[1] and Claude E. Shannon.[3] And considering it, we were always to specify at least one partition of the probability space.[2]

Thus we have designed to discuss it without discriminating the types of probability distributions.

§ 1. $\log_2 \frac{1}{P}$ and $P \log_2 \frac{1}{P}$

As the "state" (A) considered about some object is defined by the possible K cases, by the K attributes or more generally by the number K, the capability of the "source" for causing the state (A) to occur is measured by $\log_2 K$.

Mathematically this corresponds with the fact that the subsets of the finite set consisting of M elements are 2^M in all, while K and $\log_2 K$ correspond to 2^M and M respectively.

Further, if the capability is expressed statistically, the probability P in which the state (A) occurs is used for K or 2^M and it will be measured by $\log_2 \frac{1}{P}$ avoiding negative.

This quantity $\log_2 \frac{1}{P}$ will be called the amount of information for the source due to the state (A) that happens.

Hence putting conveniently $0 \log \frac{1}{0} = 0$, we have easily the following proposition.

(1.1) $P \log \frac{1}{P}$, $P \geq 0$, is concave, and attains its maximum at $P = \frac{1}{e}$, thus it is an increasing function in $0 \leq P \leq \frac{1}{e}$, ($e$ is the base of the natural logarithm); and for $\Delta P \geq 0$ and $1 \geq P_k$, $P_k + (-1)^{k+1} \Delta P \geq 0$, $k = 1, 2$,

$$(P_1 + \Delta P) \log \frac{1}{P_1 + \Delta P} - (P_2 - \Delta P) \log \frac{1}{P_2 - \Delta P} \lesseqqgtr P_1 \log \frac{1}{P_1} + P_2 \log \frac{1}{P_2}$$

accordingly as $|(P_1 + \Delta P) - (P_2 - \Delta P)| \gtreqqless |P_1 - P_2|$.

Further, as a great number of states are sometimes supposed to exist in observing some object, the average amount of information for the capability of the source due to the sequence of states such as $\{A_i \mid i = 0, 1, 2, \cdots\}$, where (A_i) is defined by the probability P_i,