The Annals of Probability
1989, Vol. 17, No. 3, 840–865

KOLMOGOROV'S CONTRIBUTIONS TO INFORMATION THEORY AND ALGORITHMIC COMPLEXITY

BY THOMAS M. COVER, PETER GACS[1] AND ROBERT M. GRAY

Stanford University, Boston University and Stanford University

1. Introduction. Kolmogorov's sustained interest in randomness and complexity led to his early contributions to information theory through seminars and papers in the 1950s, and culminated in the crucial idea of algorithmic (or descriptional) complexity in the 1960s.

Briefly, information theory says a random object $X \sim p(x)$ has complexity (entropy) $H = -\sum p(x)\log p(x)$, with the attendant interpretation that H bits are sufficient to describe X on the average. Algorithmic complexity says an object x has a complexity $K(x)$ equal to the length of the shortest (binary) program that describes x. It is a beautiful fact that these ideas are much the same. In fact, it is roughly true that $EK(X) \approx H$. Moreover, if we let $P_U(x) = \Pr\{U \text{ prints } x\}$ be the probability that a given computer U prints x when given a random program, it can be shown that $\log(1/P_U(x)) \approx K(x)$ for all x, thus establishing a vital link between the "universal" probability measure P_U and the "universal" complexity K. More on this later.

The relationship of these ideas to probability theory was summed up in Kolmogorov's 1983 paper which was based on his 1970 talk in Nice. Perhaps only the founder of modern probability theory would have the audacity to place it in the following unusual perspective ([K462][2]):

> Information theory must precede probability theory, and not be based on it. By the very essence of this discipline, the foundations of information theory have a finite combinatorial character.
>
> The applications of probability theory can be put on a uniform basis. It is always a matter of consequences of hypotheses about the impossibility of reducing in one way or another the complexity of the description of the objects in question. Naturally, this approach to the matter does not prevent the development of probability theory as a branch of mathematics being a special case of general measure theory.
>
> The concepts of information theory as applied to infinite sequences give rise to very interesting investigations, which, without being indispensable as a basis of probability theory,

Received November 1988.

[1] On leave at IBM Almaden Research Center.

[2] Reference citations preceded by K refer to the list of Kolmogorov's publications on pages 945–964.