

Combinatorial Information Distance*

Joel Ratsaby

Abstract

Let $|A|$ denote the cardinality of a finite set A . For any real number x define $t(x) = x$ if $x \geq 1$ and 1 otherwise. For any finite sets A, B let $\delta(A, B) = \log_2(t(|B \cap \bar{A}| |A|))$. We define a new combinatorial distance $d(A, B) = \max\{\delta(A, B), \delta(B, A)\}$ which may be applied to measure the distance between binary strings of different lengths. The distance is based on a classical combinatorial notion of information introduced by Kolmogorov.

Keywords: *Set distance, Lempel-Ziv complexity, Combinatorial entropy, Set entropy, Binary sequences*

*In: Advanced computational technologies, C. Enachescu, F. Gheorghe Filip, B. Iantovics (Eds.). - Bucuresti: Editura Academiei Romane (Romanian Academy Publishing House), pp. 201- 207, ISBN 978-973-27-2256-5, 2012.