Perspectives from information theory on population genetic data

Omri Tal, PhD
Max-Planck-Institut für Mathematik in den Naturwissenschaften
(Max Planck Institute for Mathematics in the Sciences)

Monday, January 23th at 12:00 pm
Jacobs Building, room no. 506

Abstract

There has recently been growing interest at borrowing both concepts and technical results from information theory for analysis in the biosciences. I will briefly review recent efforts at incorporating notions such as entropy, channel capacity, noise and distortion into biological settings. I then describe novel conceptual and quantitative links between features of population genetic samples and a core information-theoretic property. In essence, long stretches of genetic variants may be captured as ‘typical sequences’ of a nonstationary source modelled on the source population. This will provide motivation for constructing simple typicality-based population assignment schemes. I will introduce the concepts of typical genotypes, population entropy rate and mutual typicality, and their relation to the asymptotic equipartition property. Finally, I will highlight an analogy between a communication channel and an inferential channel, where ‘noise’ results from fuzzy population boundaries and parameter estimation, and where the channel capacity closely corresponds to genetic informativeness.

All Are Welcome

Contact Email:
aweber@univ.haifa.ac.il