

THE CRM APPLIED MATHEMATICAL PHYSICS (CAMP) SEMINARS



CENTRE DE RECERCA MATEMÀTICA

Ramón Ferrer i Cancho

Universitat Politècnica de Catalunya

Compression and the origins of Zipf's law of abbreviation and other linguistic laws

Abstract:

Languages across the world exhibit Zipf's law of abbreviation, namely more frequent words tend to be shorter. The generalized version of the law - an inverse relationship between the frequency of a unit and its magnitude - holds also for the behaviours of other species and the genetic code. The apparent universality of this pattern in human language and its ubiquity in other domains calls for a theoretical understanding of its origins. To this end, we generalize the information theoretic concept of mean code length as a mean energetic cost function over the probability and the magnitude of the types of the repertoire. A generalized principle of compression could explain also other linguistic laws.

Date: July 1, 2016

Place: Room C1/028

Time: 12:00

